

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-041081

(43)Date of publication of application : 08.02.2002

(51)Int.Cl.

G10L 15/18

G06F 3/16

G10L 15/06

G10L 15/00

G10L 15/28

(21)Application number : 2000-228916

(71)Applicant : SHARP CORP

(22)Date of filing : 28.07.2000

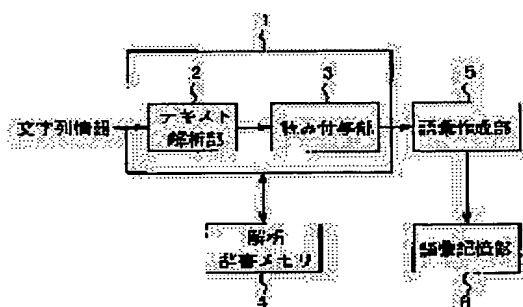
(72)Inventor : KANZA HIROYUKI
TSURUTA AKIRA

(54) UNIT/METHOD FOR PREPARING VOICE-RECOGNITION DICTIONARY, VOICE-RECOGNITION APPARATUS, PORTABLE TERMINAL, AND PROGRAM-RECORDING MEDIA

(57)Abstract:

PROBLEM TO BE SOLVED: To prepare an inexpensive voice-recognition dictionary capable of recognizing with high-accuracy.

SOLUTION: The text analyzer 2 divides the input character string into a word and finds all divided candidate words. A read impart part 3 imparts all readings to all divided candidate words. A vocabulary preparing part 5 prepares a pair of speaking unit/reading taking consideration all connection words and registers into vocabulary storage part 6. Thus the voice-recognition dictionary that can perceive correctly even with speaking of any partial character string is provided and thereby access-capable to the home page of the official residence of the prime minister with a low cost and the accuracy when the home page is entitled. 'The Top Page of the Official Residence of the Prime Minister' and made voice even as either 'The Top Page of the Official Residence of the Prime Minister', or 'The Official Residence' and or 'The Top Page of the Residence'.



LEGAL STATUS

[Date of request for examination]

26.07.2002

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the
examiner's decision of rejection or application
converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of
rejection][Date of requesting appeal against examiner's decision
of rejection]

This Page Blank (uspto)

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

This Page Blank (uspto)

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開2002-41081

(P2002-41081A)

(43)公開日 平成14年2月8日(2002.2.8)

(51)Int.Cl. ⁷	識別記号	F I	テーマコード(参考)
G 1 0 L 15/18		G 0 6 F 3/16	3 2 0 H 5 D 0 1 5
G 0 6 F 3/16	3 2 0	G 1 0 L 3/00	5 3 7 G
G 1 0 L 15/06			5 2 1 C
15/00			5 5 1 P
15/28			5 5 1 G
審査請求 未請求 請求項の数14 O L (全 25 頁)			

(21)出願番号 特願2000-228916(P2000-228916)

(22)出願日 平成12年7月28日(2000.7.28)

特許法第64条第2項ただし書の規定により図面第18図、
19図の一部は不掲載とした。

(71)出願人 000005049

シャープ株式会社

大阪府大阪市阿倍野区長池町22番22号

(72)発明者 勘座 浩幸

大阪府大阪市阿倍野区長池町22番22号 シ
ャープ株式会社内

(72)発明者 鶴田 彰

大阪府大阪市阿倍野区長池町22番22号 シ
ャープ株式会社内

(74)代理人 100062144

弁理士 青山 稔 (外1名)

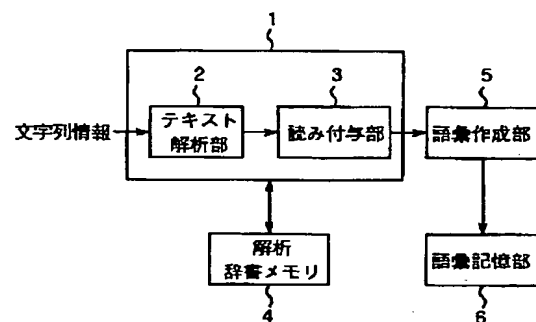
Fターム(参考) 5D015 AA04 GG01 HH01 HH04 HH11
KK02 LL05

(54)【発明の名称】 音声認識用辞書作成装置および音声認識用辞書作成方法、音声認識装置、携帯端末器、並びに、
プログラム記録媒体

(57)【要約】

【課題】 低コストで高い認識精度が得られる音声認識
用辞書を作成する。

【解決手段】 テキスト解析部2は入力文字列を単語に
分割して総ての分かち候補を求める。読み付与部3は、
総ての分かち候補の単語に対して総ての読みを付与す
る。語彙作成部5は、総ての分かち候補、総ての読み、総
ての接続単語を考慮して発声単位/読みのペアを作成し
て語彙記憶部6に登録する。こうして、1つの入力文字
列情報から、当該文字列中の何れの部分文字列を発声し
ても認識できる音声認識用辞書を作成する。したがっ
て、ホームページのタイトルが「首相官邸のトップペー
ジ」の場合に、「首相官邸のトップページ」や「首相官邸」
や「官邸のトップページ」と発声しても首相官邸のホーム
ページを呼び出すことが可能な高認識精度な音声認識用
辞書を、低コストで得ることができる。



定義 (単語単位)

後文あり?

【特許請求の範囲】

【請求項1】 入力された文字列情報を解析して構成単語に分割し、総ての分割候補を出力する解析手段と、上記分割された各構成単語に読みを付与し、総ての読み候補を出力する読み付与手段と、

上記解析手段によって得られた総ての分割候補および上記読み付与手段によって得られた総ての読み候補に基づいて、読みが付与された1つの単語または接続単語で成る1つの発声単位または単語の組み合わせが異なる複数の発声単位を認識語彙として生成する語彙作成手段と、上記生成された各認識語彙を音声認識用辞書として記憶する語彙記憶手段を備えたことを特徴とする音声認識用辞書作成装置。

【請求項2】 請求項1に記載の音声認識用辞書作成装置において、

上記解析手段は、上記分割された構成単語の列でなる各解析候補に、上記入力文字列の解析結果としての確からしさを表す解析尤度を付与するようになっており、

上記読み付与手段は、上記各解析候補を構成する単語に付与されている読みの列に、上記入力文字列の読みとしての確からしさを表す読み尤度を付与するようになっており、

上記語彙作成手段によって生成された各発声単位が存在する解析候補の上記解析尤度、上記各発声単位が存在する解析候補の上記読み尤度、上記各発声単位を構成する先頭単語の上記入力文字列中における出現順序を表す単語出現順序、上記各発声単位のモーラ数、上記各発声単位を構成する単語のうちで入力される総ての文字列情報中に最も少なく出現する単語の出現頻度を表す単語出現頻度、キーワード辞書照合結果の少なくとも一つを用いて、上記生成された各発声単位の発声確率を計算する発声確率算出手段を備えて、

上記語彙作成手段は、上記各発声単位でなる認識語彙を、上記算出された発声確率を付与して上記語彙記憶手段に記憶させるようになっていたことを特徴とする音声認識用辞書作成装置。

【請求項3】 請求項1あるいは請求項2に記載の音声認識用辞書作成装置において、

文字列情報を含むコンテンツを取り込む取り込み手段と、

辞書作成に必要な文字列情報を抽出するための抽出条件が格納された抽出条件格納手段と、

上記抽出条件を参照して、上記取り込まれたコンテンツにおける文字列情報の中から辞書作成に必要な文字列情報を抽出して上記解析手段に送出する文字列情報抽出手段を備えたことを特徴とする音声認識用辞書作成装置。

【請求項4】 請求項3に記載の音声認識用辞書作成装置において、

上記取り込み手段は、上記コンテンツとして、ウェブ・ブラウザによって表示されているウェブページの情報を

取り込むようになっていることを特徴とする音声認識用辞書作成装置。

【請求項5】 請求項3に記載の音声認識用辞書作成装置において、

上記取り込み手段は、上記コンテンツとして、文字情報化されたテレビ番組の情報を取り込むようになっていることを特徴とする音声認識用辞書作成装置。

【請求項6】 請求項2に記載の音声認識用辞書作成装置において、

上記語彙作成手段によって生成された各発声単位間の音響的な類似度を計算する類似度算出手段を備えて、

上記語彙作成手段は、上記各認識語彙に付与する発声確率を、上記算出された類似度に応じて変更するようになっていたことを特徴とする音声認識用辞書作成装置。

【請求項7】 入力された文字列情報を解析して構成単語に分割し、総ての分割候補を出力するステップと、

上記分割された各構成単語に読みを付与し、総ての読み候補を出力するステップと、

上記単語分割の結果得られた総ての分割候補および上記読み付与の結果得られた総ての読み候補に基づいて、読みが付与された1つの単語または接続単語で成る1つの発声単位または単語の組み合わせが異なる複数の発声単位を認識語彙として生成するステップと、

上記生成された各認識語彙を音声認識用辞書として記憶するステップを備えたことを特徴とする音声認識用辞書作成方法。

【請求項8】 入力された音声で、辞書に登録されている認識語彙との照合を行って認識する音声認識装置であって、

上記辞書として、請求項1乃至請求項6の何れか一つに記載の音声認識用辞書作成装置によって作成された音声認識用辞書を用いることを特徴とする音声認識装置。

【請求項9】 請求項8に記載された音声認識装置であって、

上記辞書は、請求項4に記載の音声認識用辞書作成装置によって作成された音声認識用辞書であり、

認識結果に応じたウェブページを表示するウェブページ表示手段と、

上記認識結果に基づいて、上記ウェブページ表示手段の表示内容を切換え制御する制御手段を備えたことを特徴とする音声認識装置。

【請求項10】 請求項8に記載された音声認識装置であって、

上記辞書は、請求項5に記載の音声認識用辞書作成装置によって作成された音声認識用辞書であり、

認識結果に応じたテレビ番組を表示するテレビ表示手段と、

認識結果に応じたテレビ番組を録画する録画手段と、上記録画手段によって録画されたテレビ番組を再生する再生手段と、

上記認識結果に基づいて、上記テレビ表示手段、録画手段および再生手段を制御して、表示チャンネルの切換え、録画条件の設定あるいは録画番組の再生を行う制御手段を備えたことを特徴とする音声認識装置。

【請求項11】 請求項8に記載された音声認識装置において、

特定の文字列情報の解析結果によらずに得られた認識語彙が登録された補助辞書と、

上記辞書および補助辞書との照合を行う照合手段と、
上記照合手段によって、上記認識結果として上記補助辞書に登録された認識語彙が選択された場合、上記辞書を作成する際上記音声認識用辞書作成装置に入力された文字列情報に関連する文字情報の中から当該認識結果に該当する文字列を検索する検索手段と、

上記検索された複数の文字列の中から、上記辞書に登録する文字列を選択する選択手段を備えたことを特徴とする音声認識装置。

【請求項12】 入力された音声、辞書に登録されている認識語彙との照合を行って認識する音声認識装置であって、

請求項1乃至請求項6のいずれか一つに記載の音声認識用辞書作成装置を搭載し、

上記音声認識用辞書作成装置によって作成された音声認識用辞書を、上記辞書として用いることを特徴とする音声認識装置。

【請求項13】 請求項8乃至請求項12の何れか一つに記載された音声認識装置を搭載したことを特徴とする携帯端末器。

【請求項14】 コンピュータを、
請求項1における解析手段、読み付与手段、語彙作成手段および語彙記憶手段として機能させる辞書作成処理プログラムが記録されたことを特徴とするコンピュータ読出し可能なプログラム記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】この発明は、不特定話者を対象とした音声認識装置に用いられる辞書を作成する音声認識用辞書作成装置および音声認識用辞書作成方法、作成された辞書を用いた音声認識装置、この音声認識装置を搭載した携帯端末器、並びに、辞書作成処理プログラムを記録したプログラム記録媒体に関する。

【0002】

【従来の技術】従来より、漢字仮名混じり文を含む文字列情報から音声認識用辞書を作成する方法として、上記文字列情報の形態素解析を行ってその表記の読みを求め、得られた読みを登録する方法がある。ウェブブラウザにおいて、音声認識結果に基づいてウェブページを表示するアプリケーションがあるが、その場合に使用する音声認識用辞書に、上記ウェブページを呼び出す際にユーザが発声する単語を予め登録する場合などに上記辞

書作成方法が適用される。例えば、「首相官邸」という単語を首相官邸のURL(ユニフォーム・リソース・ロケーション)と関連付けて記憶することによって、「首相官邸」とユーザが発声すれば首相官邸のホームページを表示することができる。これは、「首相官邸」という表記から「しゅしょうかんでい」という読みを求めて音声認識用辞書に登録することで実現している。

【0003】また、大量の学習用テキストデータベースから接続関係を抽出して、統計的接続情報を求めて音声認識用辞書(言語モデル)を作成する方法がある。例えば、特開平11-259088号公報に記載された音声認識装置においては、形態素解析プログラム等を用いて文字列情報を各単語(形態素)に自動的に区切り、区切られた各単語間の統計的連鎖関係、具体的にはバイグラムやトライグラム等を計算しておく。また、同じ表記の漢字であって複数の読み方がある場合には、複数の読み夫々の頻度を求めておく。そして、音声が入力されると、認識文候補を生成し、抽出された特徴パラメータを用いて言語モデルから計算される各認識文候補の尤度に、各認識文候補の単語系列を構成する各単語に上記統計的連鎖関係の確率値と読みの頻度を組み合わせて言語尤度を算出する。そして、この言語尤度に基づいて認識結果を得ている。

【0004】

【発明が解決しようとする課題】しかしながら、上記従来の音声認識用辞書作成方法においては、以下のような問題がある。すなわち、先ず、上記文字列情報の形態素解析から得られた表記の読みを登録する辞書作成方法の場合には、文字列情報と発声される接続単語を表す発声単位とが必ずしも一致するとは限らないという問題がある。以下、上述した音声認識結果を用いてブラウザによってホームページを表示する場合を例に説明する。上記ホームページの情報には、URLやタイトルがある。URLはホームページの在りかを表す情報であって、「http://www.kantei.go.jp」のごとく表記される。また、タイトルはブラウザによって決まった位置に表示されるページのタイトルであって、「首相官邸トップページ」のような漢字仮名交じりやアルファベット等の文字情報である。

【0005】上記タイトル「首相官邸トップページ」に当該辞書作成方法を適用して、上記タイトルから認識語彙を自動的に生成する場合には、タイトル「首相官邸トップページ」の形態素解析を行い、単語「首相官邸トップページ」の読みを特定することによって実現できる。

【0006】ところが、当該辞書作成方法を適用した場合には、タイトル「首相官邸トップページ」をそのまま読み「しゅしょうかんでいとつぷページ」に変換して登録するだけであるため認識語彙は「首相官邸トップページ」となり、上記タイトル「首相官邸トップページ」の部分文字列「首相官邸」を切り出して認識語彙にすることはできな

い。したがって、例えば「首相官邸」という部分発声がなされた場合には、上述のようにして作成された音声認識用辞書を用いて入力音声「首相官邸」を認識することはできないことになる。

【0007】仮に、上記ホームページのタイトルからキーワードとなる部分文字列を切り出して認識語彙を作成するにしても、部分文字列のあらゆる組み合わせや部分文字列間の発声確率等が考慮されないために、実用的な音声認識用の辞書を作成することができないという問題がある。

【0008】次に、上記統計的接続情報を求めて音声認識用辞書を作成する辞書作成方法の場合には、各単語間の区切り統計的隣接関係を用いて入力音声を認識するために上述のような問題は発生しない。また、複数の読み方がある単語の夫々の読みの頻度も含めて認識文候補の出現確率が算出される点で、より実用的な音声認識用辞書を作成することができる可能性はある。

【0009】しかしながら、単に出現確率を用いるだけでは、タイトルの前部分の表記と後ろ部分の表記とでは前者の方が発声される確率が高い(例えばタイトル「首相官邸のトップページ」の場合は「首相官邸」の方が「トップページ」よりも発声される確率が高い)といった考慮がされず、所望の認識結果が得られる確率が低いという問題がある。

【0010】また、「ホームページ」や「ようこそ」等の特定の語彙は単独では発声されないと考えられる。ところが、それに対する考慮もされていないため、別の発声が「ホームページ」という語彙に誤認識されてしまう可能性がある。

【0011】さらに、このような大量の学習テキストデータベースから音声認識用辞書(言語モデル)を作成する辞書作成方法は、ディクテーションのような数万語の語彙を認識する用途にはよいが、高速のCPU(中央演算処理装置)と大量の記憶容量が必要であるというコスト上の問題と、非常に多数の語彙の中から正しいものを見つけ出すのは困難であるという認識性能上の問題とがある。

【0012】そこで、この発明の目的は、低コストで高い認識精度が得られる音声認識用辞書作成装置および音声認識用辞書作成方法、作成された辞書を用いた音声認識装置、この音声認識装置を搭載した携帯端末器、並びに、辞書作成処理プログラムを記録したプログラム記録媒体を提供することにある。

【0013】

【課題を解決するための手段】上記目的を達成するため、第1の発明の音声認識用辞書作成装置は、入力された文字列情報を解析して構成単語に分割し、総ての分割候補を出力する解析手段と、上記分割された各構成単語に読みを付与し、総ての読み候補を出力する読み付与手段と、上記解析手段によって得られた総ての分割候補お

よび上記読み付与手段によって得られた総ての読み候補に基づいて、読みが付与された1つの単語または接続単語で成る1つの発声単位または単語の組み合わせが異なる複数の発声単位を認識語彙として生成する語彙作成手段と、上記生成された各認識語彙を音声認識用辞書として記憶する語彙記憶手段を備えたことを特徴としている。

【0014】上記構成によれば、一つの文字列情報から得られた総ての分割候補および総ての読み候補に基づいて、語彙作成手段によって、読みが付与された1つ又は複数の発声単位が生成される。したがって、こうして生成された発声単位を認識語彙として登録することによって、与えられた文字列情報から、発声の可能性がある発声単位を認識語彙とする音声認識用辞書が生成される。すなわち、利用者が、予め設定された文字列中のどの部分文字列を発声しても正しく認識できる音声認識装置を実現可能な音声認識用辞書が作成されるのである。

【0015】また、上記第1の発明の音声認識用辞書作成装置は、上記解析手段を、上記分割された構成単語の列でなる各解析候補に、上記入力文字列の解析結果としての確からしさを表す解析尤度を付与するように成し、上記読み付与手段を、上記各解析候補を構成する単語に付与されている読みの列に、上記入力文字列の読みとしての確からしさを表す読み尤度を付与するように成し、上記語彙作成手段によって生成された各発声単位が存在する解析候補の上記解析尤度、上記各発声単位が存在する解析候補の上記読み尤度、上記各発声単位を構成する先頭単語の上記入力文字列中における出現順序を表す単語出現順序、上記各発声単位のモーラ数、上記各発声単位を構成する単語のうちで入力される総ての文字列情報中に最も少なく出現する単語の出現頻度を表す単語出現頻度、キーワード辞書照合結果の少なくとも一つを用いて、上記生成された各発声単位の発声確率を計算する発声確率算出手段を備えて、上記語彙作成手段を、上記各発声単位でなる認識語彙を、上記算出された発声確率を付与して上記語彙記憶手段に記憶させるように成すことが望ましい。

【0016】上記構成によれば、上記語彙記憶手段には、上記各発声単位でなる認識語彙に解析尤度、読み尤度、単語出現順序、モーラ数、単語出現頻度およびキーワード辞書照合結果の少なくとも一つを用いて算出された発声確率が付与されて登録されている。したがって、上記音声認識用辞書中に、上記解析手段による誤解析の結果生成された認識語彙や発声されない認識語彙が登録されていても、このような不要な認識語彙の発声確率が小さく設定されて、高い認識精度を呈する音声認識装置を実現できる音声認識用辞書が作成される。

【0017】また、上記第1の発明の音声認識用辞書作成装置は、文字列情報を含むコンテンツを取り込む取り込み手段と、辞書作成に必要な文字列情報を抽出するた

めの抽出条件が格納された抽出条件格納手段と、上記抽出条件を参照して、上記取り込まれたコンテンツにおける文字列情報の中から辞書作成に必要な文字列情報を抽出して上記解析手段に送出する文字列情報抽出手段を備えることが望ましい。

【0018】上記構成によれば、抽出条件格納手段にコンテンツの特徴を利用した抽出条件を格納しておくことによって、コンテンツ情報から自動的に上記音声認識用辞書が作成される。

【0019】また、上記第1の発明の音声認識用辞書作成装置は、上記取り込み手段を、上記コンテンツとして、ウェブ・ブラウザによって表示されているウェブページの情報を取り込むように成すことが望ましい。

【0020】上記構成によれば、上記抽出条件格納手段に、例えば、「<title>タグが存在する場合は<title>と</title>とで囲まれた文字列を抽出する」を格納しておくことによって、ウェブページのタイトル「<title>首相官邸のトップページ</title>」から文字列「首相官邸のトップページ」が抽出される。そして、上記文字列「首相官邸のトップページ」に基づいて、上述のようにして自動的に音声認識用辞書が作成される。

【0021】また、上記第1の発明の音声認識用辞書作成装置は、上記取り込み手段を、上記コンテンツとして、文字情報化されたテレビ番組の情報を取り込むように成すことが望ましい。

【0022】上記構成によれば、上記抽出条件格納手段に、例えば、「番組名というタグが付いている文字列を抽出する」を格納しておくことによって、タグ「番組名」に該当する文字列「NHKニュースおはよう日本」が抽出される。そして、上記文字列「NHKニュースおはよう日本」に基づいて、上述のようにして自動的に音声認識用辞書が作成される。

【0023】また、上記第1の発明の音声認識用辞書作成装置は、上記語彙作成手段によって生成された各発声単位間の音響的な類似度を計算する類似度算出手段を備え、上記語彙作成手段は、上記各認識語彙に付与する発声確率を上記算出された類似度に応じて変更するように成すことが望ましい。

【0024】上記構成によれば、上記語彙作成手段によって生成された発声単位「首相(しゅしょう)」と発声単位「主張(しゅちょう)」との類似度が所定値よりも高く、両発声単位が音響的に類似している場合には、例えば、発声確率の値が高く入力文字列中において中心的な役割を果たす発声単位「首相」の発声確率の値が更に高められる一方、そうでない発声単位「主張」の発声確率の値が更に低められる。こうすることによって、中心的な役割を果たす発声「首相」が「主張」と誤認識されて、例えば、目的とする「首相官邸のトップページ」のホームページではなく、「司法省の主張」のホームページが表示されてしまうことが防止される。

【0025】また、第2の発明の音声認識用辞書作成方法は、入力された文字列情報を解析して構成単語に分割し、総ての分割候補を出力するステップと、上記分割された各構成単語に読みを付与し、総ての読み候補を出力するステップと、上記単語分割の結果得られた総ての分割候補および上記読み付与の結果得られた総ての読み候補に基づいて、読みが付与された1つの単語または接続単語で成る1つの発声単位または単語の組み合わせが異なる複数の発声単位を認識語彙として生成するステップと、上記生成された各認識語彙を音声認識用辞書として記憶するステップを備えたことを特徴としている。

【0026】上記構成によれば、上記第1の発明の場合と同様に、読みが付与された1つまたは複数の発声単位が生成される。したがって、こうして生成された発声単位を認識語彙として登録することによって、与えられた文字列情報から、発声の可能性のある発声単位を認識語彙とする音声認識用辞書が生成される。すなわち、予め設定された文字列の部分文字列を発声しても正しく認識できる音声認識装置を実現可能な音声認識用辞書が作成されるのである。

【0027】また、第3の発明は、入力された音声を、辞書に登録されている認識語彙との照合を行って認識する音声認識装置であって、上記辞書として、上記第1の発明の音声認識用辞書作成装置によって作成された音声認識用辞書を用いることを特徴としている。

【0028】上記構成によれば、与えられた文字列情報から生成された発声の可能性のある発声単位を認識語彙とする音声認識用辞書との照合を行って、入力音声は認識される。したがって、予め設定された文字列の部分文字列を発声しても正しく認識される。

【0029】また、上記第3の発明の音声認識装置は、上記辞書を、上記コンテンツとしてウェブページの情報を取り込む音声認識用辞書作成装置によって作成された音声認識用辞書と成し、認識結果に応じたウェブページを表示するウェブページ表示手段と、上記認識結果に基づいて、上記ウェブページ表示手段の表示内容を切換え制御する制御手段を備えることが望ましい。

【0030】上記構成によれば、上記ウェブページの情報から自動的に作成された音声認識用辞書が用いられているので、ウェブページのタイトル等が正しく認識される。したがって、制御手段によって、認識結果に基づいてウェブページ表示手段の表示内容が切換え制御されることによって、上記ウェブページ表示手段に認識結果に応じたウェブページが正しく表示される。

【0031】また、上記第3の発明の音声認識装置は、上記辞書を、上記コンテンツとしてテレビ番組の情報を取り込む音声認識用辞書作成装置によって作成された音声認識用辞書と成し、認識結果に応じたテレビ番組を表示するテレビ表示手段と、認識結果に応じたテレビ番組を録画する録画手段と、上記録画手段によって録画され

たテレビ番組を再生する再生手段と、上記認識結果に基づいて、上記テレビ表示手段・録画手段および再生手段を制御して、表示チャンネルの切換え・録画条件の設定あるいは録画番組の再生を行う制御手段を備えることが望ましい。

【0032】上記構成によれば、上記テレビ番組の情報から自動的に作成された音声認識用辞書が用いられているので、テレビ番組名等が正確に認識される。したがって、制御手段によって、認識結果に基づいてテレビ表示手段、録画手段および再生手段が制御されることによって、表示チャンネルの切換え、録画条件の設定あるいは録画番組の再生が正しく行われる。

【0033】また、上記第3の発明の音声認識装置は、特定の文字列情報の解析結果によらずに得られた認識語彙が登録された補助辞書と、上記辞書および補助辞書との照合を行う照合手段と、上記照合手段によって、上記認識結果として上記補助辞書に登録された認識語彙が選択された場合、上記辞書を作成する際に上記音声認識用辞書作成装置に入力された文字列情報に関連する文字情報の中から当該認識結果に該当する文字列を検索する検索手段と、上記検索された複数の文字列の中から、上記辞書に登録する文字列を選択する選択手段を備えることが望ましい。

【0034】上記構成によれば、上記第1の発明の音声認識用辞書作成装置によって作成された辞書に登録されていない語彙を発声した場合でも、その語彙が正しく認識される。さらに、認識結果として、上記辞書に登録された認識語彙ではなく、上記第1の発明の音声認識用辞書作成装置によらずに作成された補助辞書に登録された認識語彙が選択された場合には、検索手段によって、上記音声認識用辞書作成装置に入力された例えばウェブページのタイトルに関連するウェブページ情報の中から、当該認識結果に該当する文字列が検索される。そして、選択手段によって、上記検索された複数の文字列の中から上記辞書に登録する文字列が選択される。したがって、その語彙を上記辞書に登録することによって、上記辞書の認識語彙数が増加して認識速度が向上される。

【0035】また、第4の発明は、入力された音声、辞書に登録されている認識語彙との照合を行って認識する音声認識装置であって、上記第1の発明の音声認識用辞書作成装置を搭載し、上記音声認識用辞書作成装置によって作成された音声認識用辞書を、上記辞書として用いることを特徴としている。

【0036】上記構成によれば、搭載されている音声認識用辞書作成装置に文字列情報を入力することによって、この文字列情報から発声の可能性がある発声単位が生成され、この発声単位を認識語彙とする音声認識用辞書が作成される。したがって、この音声認識用辞書との照合を行って入力音声を認識することによって、予め設定された文字列の部分文字列を発声しても正しく認識さ

れる。

【0037】また、第5の発明の携帯端末器は、上記第3、第4の発明の音声認識装置を搭載したことを特徴としている。

【0038】携帯端末器においては、操作指示を行う場合にはキー操作よりも発声による方が操作性はよい。上記構成によれば、予め設定された文字列中のどの部分文字列を発声しても正しく認識できる音声認識用辞書を用いた音声認識装置が搭載されている。したがって、出先等において操作指示を行うための文言を予め決められている通りに正確に発声しなくとも、例えばホームページの呼び出し等の操作が正しく行われる。

【0039】また、第6の発明のプログラム記録媒体は、コンピュータを、上記第1の発明における解析手段、読み付与手段、語彙作成手段および語彙記憶手段として機能させる辞書作成処理プログラムが記録されていることを特徴としている。

【0040】上記構成によれば、上記第1の発明の場合と同様に、読みが付与された1つまたは複数の発声単位が生成される。したがって、こうして生成された発声単位を認識語彙として登録することによって、予め設定された文字列の部分文字列を発声しても正しく認識できる音声認識装置を実現可能な音声認識用辞書が作成されるのである。

【0041】

【発明の実施の形態】以下、この発明を図示の実施の形態により詳細に説明する。

<第1実施の形態>図1は、本実施の形態の音声認識用辞書作成装置におけるブロック図である。解析処理部1に文字列情報が入力されると、テキスト解析部2によって入力文字列の言語が解析されて形態素に分割される。その際に、複数の分割候補が存在する場合は、その総ての分割候補が出力される。そして、読み付与部3によって、上記分割された形態素の読みが付与される。その際に、複数の読み方が存在する場合には、その総ての読みが出力される。解析辞書メモリ4には、上記テキスト解析部2がテキスト解析を行う際に必要な解析辞書を含む言語データ等が格納されている。

【0042】語彙作成部5は、上記テキスト解析部2によるテキスト解析結果と読み付与部3による読み付与結果を元に、音声認識を行うために必要な音声認識用辞書を作成する。語彙記憶部6は、語彙作成部5によって作成された音声認識用辞書を記憶する。そして、この音声認識用辞書は音声認識時に使用される。

【0043】図2は、上記構成を有する音声認識用辞書作成装置の各部によって実行される辞書作成処理動作のフローチャートである。以下、図2のフローチャートに従って、解析処理部1に「浦和市大久保」という文字列情報が入力された場合を例に、本音声認識用辞書作成装置の動作を説明する。ここで、テキスト解析部2に対する

文字列情報(テキスト)の入力は、WWW(World Wide Web)等のネットワークからの入力や受信された文字放送による入力であってもよいし、キーボードやペン等の文入力手段からの入力あるいは音声認識装置からの認識結果の入力でもよいし、OCR(光学式文字読取り装置)等の文字認識装置からの入力であっても構わない。

【0044】ステップS1で、上記テキスト解析部2によって文字列情報「浦和市大久保」が取り込まれる。ステップS2で、さらに、入力文字列「浦和市大久保」が、解析辞書メモリ4に格納されている解析辞書を参照して形態素(単語)に分割されるテキスト解析が行われる。その結果、単語列「浦和(名詞)」、「市(接尾語)」、「大久保(名詞)」と、単語列「浦和(名詞)」、「市大(名詞)」、「久保(名詞)」との2つの分かち候補が得られる。尚、図3に、テキスト解析部2の出力結果の一例を示す。

【0045】ステップS3で、上記読み付与部3によって、各単語に読みが付与されて、図4に示すような読み付与結果が得られる。ここで、各単語の読みは、解析辞書メモリ4に格納されている解析辞書を参照して付与される。尚、複数の読み方が存在する単語には総ての読みが付与される。図4においては、文字「市」には、読み「し」と読み「いち」との2つの読みが付与されている。以下、語彙作成部5に処理が移行する。

【0046】ステップS4で、上記テキスト解析部2によって得られた分かち候補の数を変数 k にセットされる。上記入力文字列「浦和市大久保」の場合は、図3に示すように「浦和/市/大久保」と「浦和/市大/久保」との2通り分かち候補が存在するので、変数 k に「2」がセットされる。ステップS5で、分かち候補番号 m に初期値「1」がセットされる。ステップS6で、 m 番目の分かち候補を構成する総単語数を変数 N_m にセットされる。上記入力文字列「浦和市大久保」の場合は、1番目の分かち候補「浦和/市/大久保」は、「浦和(名詞)」、「市(接尾語)」および「大久保(名詞)」の3単語であるから、変数 N_1 に「3」がセットされる。

【0047】ステップS7で、語彙として登録する発声単位を構成する二つ目以降の単語数 j が、「0」に初期化される。ステップS8で、発声単位的位置(先頭単語の番号) i が、「1」に初期化される。ステップS9で、 $(i+j)$ の値が総単語数 N_m 以下であるか否かが判別される。その結果、総単語数 N_m 以下である場合にはステップS10に進む一方、総単語数 N_m より大きい場合にはステップS13に進む。ステップS10で、発声単位を構成する単語 W_i, \dots, W_{i+j} と、対応する読み $Y_i(1), \dots, Y_i(1+p_i), \dots, Y_{i+j}(1), \dots, Y_{i+j}(1+p_{i+j})$ (但し、 p_{i+j} : 単語 W_{i+j} の二つ目以降の読みの数)とが、語彙記憶部6に登録される。したがって、1つの単語 W_i に対する読みが $Y_i(1), Y_i(2), \dots$ のごとく複数ある場合には総ての読みが登録される。

【0048】ステップS11で、上記発声単位的位置 i の

値がインクリメントされる。ステップS12で、発声単位的位置 i が総単語数 N_m 以下であるか否かが判別される。その結果、総単語数 N_m 以下である場合にはステップS9に戻って次の位置に在る発声単位の登録に移行する。一方、総単語数 N_m より大きい場合にはステップS13に進む。ステップS13で、発声単位を構成する二つ目以降の単語数 j がインクリメントされる。ステップS14で、上記単語数 j が総単語数 N_m より小さいか否かが判別される。その結果、総単語数 N_m より小さい場合にはステップS8に戻って二つ目以降の単語数が1つ多い発声単位の登録に移行する。一方、単語数 N_m 以上である場合にはステップS15に進む。ステップS15で、分かち候補番号 m がインクリメントされる。ステップS16で、分かち候補番号 m が分かち候補の数 k 以下であるか否かが判別される。その結果、分かち候補数 k 以下である場合にはステップS6に戻って次の分かち候補に対する処理に移行する。一方、分かち候補数 k より大きい場合には、総ての分かち候補に関する処理が終了したと判断されて辞書作成処理動作を終了する。

【0049】以上の結果、上記文字列情報「浦和市大久保」に対して上述のような辞書作成処理動作が行われた場合には、語彙記憶部6へは次のように登録が行われる。すなわち、1番目の分かち候補「浦和/市/大久保」に対して、先ず、発声単位の数二つ目以降の単語数 j が「0」である場合には、発声単位的位置 i が小さい順に、発声単位/読みのペア「浦和/うらわ」、「市/し」、「大久保/おおくぼ」が登録される。次に、発声単位の上記単語数 j が「1」である場合は、発声単位的位置 i が小さい順に、発声単位/読みのペア「浦和市/うらわし」、「市大久保/しおおくぼ」が登録される。次に、発声単位の上記単語数 j が「2」である場合には、発声単位/読みのペア「浦和市大久保/うらわしおおくぼ」が登録される。

【0050】さらに、2番目の分かち候補「浦和/市大/久保」に対して、先ず、発声単位の上記単語数 j が「0」である場合は、発声単位的位置 i が小さい順に、発声単位/読みのペア「浦和/うらわ」、「市大/しだい」、「久保/くぼ」が登録される。その際に、他の発声単位/読みのペア「市大/いちだい」も登録される。次に、発声単位の上記単語数 j が「1」である場合には、発声単位/読みのペア「浦和市大/うらわしだい」、「市大久保/しだいくぼ」と、他の発声単位/読みのペア「浦和市大/うらわいちだい」、「市大久保/いちだいくぼ」とが登録される。次に、発声単位の上記単語数 j が「2」である場合には、発声単位/読みのペア「浦和市大久保/うらわしだいくぼ」と、他の発声単位/読みのペア「浦和市大久保/うらわいちだいくぼ」とが登録される。

【0051】その結果、上記語彙記憶部6には、図5に示すように認識語彙が登録されて音声認識用辞書が作成されることになる。

【0052】以上のように、本実施の形態においては、

テキスト解析部2によって、入力文字列が形態素(単語)に分割されて、総ての分かち候補が求められる。さらに、読み付与部3によって、総ての分かち候補の単語に対して読みが付与される。その場合、複数の読み方が存在する単語には総ての読みが付与される。そして、語彙作成部5によって、総ての分かち候補、総ての読み候補、総ての接続単語の組み合わせを考慮して発声単位/読みのペアを作成し、語彙記憶部6に登録するようにしている。

【0053】すなわち、本実施の形態によれば、入力文字列情報から複数の発声単位の語彙を生成することができる。したがって、1つの入力文字列情報から、当該文字列中の何れの部分文字列を発声しても認識できる音声認識用辞書を作成することが可能となるのである。

【0054】例えば、ウェブ上でホームページの呼び出しを音声によって行う場合に、テキスト解析部2にホームページのタイトルを入力し、上述の手順によって音声認識用辞書を作成することによって、上記タイトルと発声が完全に一致しなくてもホームページの呼び出しを行うことができるのである。例えば、ホームページのタイトルが「首相官邸のトップページ」である場合、「首相官邸のトップページ」と発声しても、「首相官邸」と発声しても、あるいは、「官邸のトップページ」と発声しても、首相官邸のホームページを呼び出すことが可能な音声認識用辞書を得ることができるのである。

【0055】同様に、テレビのチャンネル切替をテレビ番組名の発声によって行う場合に、テキスト解析部2に電子テレビ番組表の番組名を入力し、上述の手順によって音声認識用辞書を作成することによって、例えば、番組名が「NHKニュースおはよう日本」である場合に、「NHKニュース」と発声しても、「おはよう日本」と発声しても、所定の時間になるとテレビのチャンネルを自動的に「NHK」に切り換えることが可能な音声認識用辞書を得ることができるのである。

【0056】尚、説明を簡単にするため、文字列「NHKニュース」、「おはよう日本」を一つの単語として音声認識用辞書に登録する場合を例に説明を行っている。しかしながら、「NHK」と「ニュース」や「おはよう」と「日本」を独立した単語として記憶して、これらの単語が連続するという情報を別途記憶することによって、連続音声認識用の辞書作成も同様の方法で実現することができる。

【0057】＜第2実施の形態＞上記第1実施の形態によれば、テキスト解析の結果得られる接続単語の組合せだけの発声単位が得られるが、これらの発声単位の中には誤解析も含まれる。したがって、得られた発声単位の総てを同等に扱うことは不要な語彙を多く含み、認識率を低下させる原因になり得る。本実施の形態は、このような場合に対処するものである。

【0058】図6は、本実施の形態の音声認識用辞書作成装置におけるブロック図である。解析処理部11に文

字列情報が入力されると、テキスト解析部12によって入力文字列が形態素に分割される。その際に、複数の分割候補が存在する場合には、その総ての分割候補に確からしさの度合いを表す解析尤度が付与されて出力される。そして、読み付与部13によって、上記分割された形態素の読みが付与される。その際に、複数の読み方が存在する場合には、その総ての読みに確からしさの度合いを表す読み尤度が付与されて出力される。解析辞書メモリ14には、上記第1実施の形態の場合と同様に、テキスト解析部12がテキスト解析を行う際や読み付与部13が読み付与を行う際に必要な解析辞書を含む言語データ等が格納されている。

【0059】語彙作成部15は、テキスト解析部12によるテキスト解析結果と、読み付与部13による読み付与結果と、発声確率計算部16による発声確率とを基に、音声認識を行うために必要な音声認識用辞書を作成する。発声確率計算部16は、テキスト解析部12が出力する解析尤度、読み付与部13が出力する読み尤度、単語出現順序、モーラ数、単語出現頻度、キーワード辞書照合の少なくとも一つを用いて、語彙作成部15によって各分かち候補から得られた発声単位の発声確率を算出する。語彙記憶部17は、語彙作成部15によって作成された音声認識用辞書を記憶する。そして、この音声認識用辞書は音声認識時に使用される。

【0060】図7に、テキスト解析部12の出力結果の一例を示す。また、図8に、読み付与部13の出力結果の一例を示す。また、図9に、語彙作成部15の出力結果の一例を示す。

【0061】以下、本実施の形態における特徴である発声確率計算部16の動作について詳細に述べる。この発声確率計算部16では、上述したように、テキスト解析部12による解析尤度KS、読み付与部13による読み尤度YS、単語出現順序、モーラ数、単語出現頻度、キーワード辞書照合の少なくとも一つを用いて、各発声単位の発声確率を算出する。

【0062】先ず、上記テキスト解析部12によって得られる解析尤度KSから上記発声確率を求める場合について説明する。上記解析尤度KSは、入力文字列を解析して得られた結果(形態素分割結果)がどの程度確からしさを図る指標になる。上記第1実施の形態の場合と同じ入力文字列「浦和市大久保」を例に上げて説明する。

【0063】上記テキスト解析部12によって、1番目の解析候補(分かち候補)「浦和/市/大久保」の解析尤度がKS(1)と得られ、2番目の解析候補「浦和/市大/久保」の解析尤度がKS(2)と得られたとする。また、i番目の解析候補に含まれる構成単語を組み合わせて得られる発声単位の数をN(i)とする。但し、同じ発声単位が複数の解析候補に含まれている場合には、最も解析尤度の高い解析候補の発声単位を数えることとする。上記入力文字列「浦和市大久保」の場合であれば、構成単語が「浦

和」だけで成る発声単位は、1番目、2番目の両解析候補に含まれるために、解析尤度の高い解析候補の方だけで数えるのである。

【0064】一つの解析候補における構成単語の全組み

$$P1(w) = KS(i) / \sum_{j=1}^M (KS(j) \times N(j)) \quad \dots(1)$$

【0065】式(1)における右辺の分母は、入力文字列に固有の値であるから「A」と置くことができる。また、一般的に、入力文字列「浦和市大久保」は、固有名詞「浦和」と接尾語「市」と固有名詞「大久保」である可能性の方が、固有名詞「浦和」と一般名詞「市大」と固有名詞「久保」である可能性よりも高いと判断され、両解析尤度の大小関係は $KS(1) > KS(2)$ となる。したがって、「浦和市」と発声される確率

$KS(1)/A$

と、「久保」と発声される確率

$KS(2)/A$

とでは、前者の方が高くなる。

【0066】以上のようにして、各解析候補の解析尤度 KS に基づいて、各解析候補から得られる発声単位の発声確率を求めることができるのである。

【0067】次に、上記読み付与部13によって得られる読み尤度 YS から上記発声確率を求める場合について説明する。上記入力文字列「浦和市大久保」の場合を例に

$$P2(w) = YS(i) / \sum_{j=1}^M (YS(j) \times N(j)) \quad \dots(2)$$

【0070】式(2)における右辺の分母は、入力文字列に固有の値であるから「B」と置くことができる。また、各読み尤度の大小関係は $YS(1) > YS(2) > YS(3)$ であるとする。そうすると、「うらわしおおくぼ」と発声される確率

$YS(1)/B$

と「うらわしだいこぼ」と発声される確率

$YS(2)/B$

とでは、前者の方が高くなる。

【0071】以上のようにして、各解析候補の読み尤度 YS に基づいて、各解析候補から得られる発声単位の発声確率を求めることができるのである。

【0072】次に、上記単語出現順序から上記発声確率を求める場合について説明する。ここで、単語出現順序 i を変数とする関数 $h(i)$ を定義する。関数 $h(i)$ は、文字列表記の先頭部分にある単語ほど発声単位となる確率が高いという経験則に基づいて、単語出現順序 i が増加するに従って値が減少する関数である。

【0073】例えば電子番組データの番組名に記載されている文字列「水戸黄門/五十三次世直し旅・品川」を例に説明する。単語出現順序が i 番目である単語を先頭を含む発声単位の個数を $N(i)$ とする。

【0074】同じ単語を先頭を含む全発声単位は均等の確率で発生するものとする。すなわち、単語「水戸」で始

合わせ(つまり全発声単位)は、均等の確率で発生するものとする。そうすると、 i 番目の解析候補における発声単位 w の発声確率 $P1(w)$ は式(1)で表現できる。但し、式(1)中における M は解析候補数である。

上げて説明する。上記解析尤度が付与された総ての解析候補に対して、読み付与部13によって、1番目の解析候補「浦和(うらわ)/市(し)/大久保(おおくぼ)」の読み尤度が $YS(1)$ 、2番目の解析候補「浦和(うらわ)/市大(しだい)/久保(くぼ)」の読み尤度が $YS(2)$ 、3番目の解析候補「浦和(うらわ)/市大(いちだい)/久保(くぼ)」の読み尤度が $YS(3)$ と計算されたとする。

【0068】また、 i 番目の解析候補に含まれる構成単語を組み合わせて得られる発声単位の数を $N(i)$ とする。但し、同じ発声単位が複数の解析候補に含まれている場合には、最も解析尤度の高い解析候補の発声単位を数えることとする。

【0069】一つの解析候補における構成単語の全組合せは、均等の確率で発生するものとする。そうすると、 i 番目の解析候補における発声単位 w の発声確率 $P2(w)$ は式(2)で表現できる。但し、式(2)中における M は解析候補数である。

まる発声単位「水戸」、「水戸黄門」、「水戸黄門五十三次」の発声確率は総て等確率になる。そうすると、単語出現順序が i 番目である単語を先頭を含む発声単位 w の発声確率 $P3(w)$ は式(3)で表現できる。但し、式(3)中における M は単語出現順序数(単語数)である。

$$P3(w) = h(i) / \sum_{j=1}^M (h(j) \times N(j)) \quad \dots(3)$$

【0075】式(3)における右辺の分母は、解析候補に固有の値であるから「C」と置くことができる。また、上記文字列「水戸黄門/五十三次世直し旅・品川」に対するテキスト解析部12によるテキスト解析の結果、単語出現順序 $i=1$ は「水戸」、 $i=2$ は「黄門」、 $i=3$ は「五十三次」、 $i=4$ は「世直し」、 $i=5$ は「旅」、 $i=6$ は「品川」となる。さらに、関数 $h(i)$ の定義に基づいて、単語出現順序 $i=1, 3, 6$ における関数 $h(i)$ の値の大小関係は $h(1) > h(3) > h(6)$ となる。そのために、発声単位「水戸黄門」、「五十三次世直し旅」、「品川」に関して、発声単位「品川」よりも発声単位「五十三次世直し旅」、発声単位「五十三次世直し旅」よりも発声単位「水戸黄門」の方が、発声確率は高くなる。したがって、「みとこうもん」と発声される確率

$h(1)/C$

と「ごじゅうさんつぎ」と発声される確率

$h(3)/C$

と「しながわ」と発声される確率

$$h(6)/C$$

とでは、「みとこうもん」の発声確率が一番高くなる。

【0076】以上のようにして、入力文字列を構成する各単語の単語出現順序に基づいて、得られる発声単位の発声確率を求めることができるのである。

【0077】次に、上記モーラ数から上記発声確率を求める場合について説明する。発声単位の発声確率は、あるモーラ数をピークとして発声単位のモーラ数が大きくなるほど低くなる。逆に、モーラ数が小さくなっても低くなる。そこで、図10に概念図を示すような発声単位のモーラ数 i と発声確率との関係を表す関数 $m(i)$ を定義する。

【0078】電子番組データの番組名に記載されている文字列「サスペンス傑作劇場津軽竜飛岬風の殺意」を例に説明する。この番組名を発声する際に「サスペンス傑作劇場津軽竜飛岬風の殺意」とそのまま発声するケースは稀であり、「サスペンス傑作劇場」や「サスペンス」あるいは「津軽竜飛岬風の殺意」等の発声単位で発声される可能性が高いと考えられる。また、「傑作」や「劇場」ではモーラ数が少な過ぎて発声確率は低いと考えられる。

【0079】同じモーラ数を呈する全発声単位は均等の確率で発生するものとする。また、モーラ数が i である発声単位の個数を $N(i)$ とする。そうすると、モーラ数が i である発声単位 w の発声確率 $P_4(w)$ は式(4)で表現できる。但し、式(4)中における M は最大モーラ数である。

$$P_4(w) = m(i) / \sum_{j=1}^M (m(j) \times N(j)) \quad \dots (4)$$

【0080】式(4)における右辺の分母は、解析候補に固有の値であるから「D」と置くことができる。また、上記文字列「サスペンス傑作劇場津軽竜飛岬風の殺意」に対するテキスト解析部12によるテキスト解析の結果、発声単位「サスペンス傑作劇場津軽竜飛岬風の殺意」のモーラ数は24、発声単位「サスペンス劇場」のモーラ数は9となる。さらに、関数 $m(i)$ の定義に基づいて、モーラ数 $i = 24, 9$ における関数 $m(i)$ の値の大小関係は $m(9) > m(24)$ となる。したがって、「さすべんすげきじょうつがるたっぴみさきかぜのさつい」と発声される確率

$$m(24)/D$$

と「さすべんすげきじょう」と発声される確率

$$m(9)/D$$

とでは、後者の方が高くなる。

【0081】以上のようにして、入力文字列から得られる発声単位のモーラ数に基づいて、発声単位の発声確率を求めることができるのである。

【0082】次に、上記単語出現頻度から上記発声確率を求める場合について説明する。ここで、テキスト解析部12に順次入力される総ての文字列情報で成る文字列

情報群中における出現回数 i を変数とする関数 $f(i)$ を定義する。この関数 $f(i)$ は、上記文字列情報群の中に出現する回数が多い単語は他の単語と区別するための単語となり難いという性質を利用しており、上記文字列情報群中における出現回数 i が増加するに従って値が減少する関数である。

【0083】例えば、ホームページのタイトル「大蔵省ホームページ」を例に説明する。単語「ホームページ」や単語「ようこそ」等は、他のホームページのタイトルにも頻繁に出現するので、ホームページの呼び出し用の認識語彙としてこれらの単語を含む発声単位の発声確率は低くなる。

【0084】上記文字列情報群中において同じ出現回数を呈する単語を含む全発声単位は均等の確率で発生するものとする。また、上記文字列情報群中での出現回数が i である単語を含む発声単位の個数を $N(i)$ とする。そうすると、発声単位 w を構成する単語のうちで上記文字列情報群中での出現回数が最少である単語の出現回数が i である当該発声単位 w の発声確率 $P_5(w)$ は式(5)で表現できる。但し、式(5)中における M は最大出現回数である。

$$P_5(w) = f(i) / \sum_{j=1}^M (f(j) \times N(j)) \quad \dots (5)$$

【0085】式(5)における右辺の分母は、上記文字列情報群に固有の値であるから「E」と置くことができる。また、例えば、上記文字列情報群中に、「大蔵省」という単語を含む発声単位が1回含まれる一方、「ホームページ」という単語を含む発声単位が5回含まれるとすると、関数 $f(i)$ の定義に基づく単語出現頻度 $i = 1, 5$ における関数 $f(i)$ の値の大小関係は $f(1) > f(5)$ となる。したがって、「おおくらしょう」と発声される確率 $f(1)/E$

と「ほーむぺーじ」と発声される確率

$$f(5)/E$$

とでは、前者の方が高くなる。

【0086】以上のようにして、上記文字列情報群中における単語出現頻度に基づいて、発声単位の発声確率を求めることができるのである。

【0087】具体的には、例えば、上記文字列情報群として5つのホームページのタイトルがある場合、先ず、1番目のタイトルが入力された場合に、当該タイトルの文字列中における単語出現頻度 i_1 に基づいて発声単位 W の発声確率 $P_5(w)$ を求める。次に、2番目のタイトルが入力された場合は、同一発声単位 W に関して、1番目のタイトルと2番目のタイトルとの文字列中における単語出現頻度 i_2 に基づいて、発声確率 $P_5(w)$ を計算し直す。以下同じ動作を繰り返し、最後に5番目のタイトルが入力された場合は、同一発声単位 W に関して、1番目～5番目のタイトルの全文字列中における単語出現頻度 i_t に基づいて、発声確率 $P_5(w)$ を計算し直し、

最終的な発声単位Wの発声確率 $P_5(w)$ を得るのである。

【0088】最後に、上記キーワード辞書照合から上記発声確率を求める場合について説明する。この場合、予め確率値が付与されたキーワードを登録したキーワード辞書を作成しておく。例えば、単語「ニュース」には高い確率値を付与してキーワード辞書に登録しておく。逆に、単語「番組」や単語「ホームページ」は冗長で在っても無くてもよく、低い確率値を付与して登録しておく。尚、キーワード辞書に登録されていない単語にはデフォルトの確率値が与えられる。

【0089】こうすることによって、「にゅーす」と発声される確率と「ばんぐみ」と発声される確率とでは、前者の方が高くなる。ここで、確率値「0」を付与してキーワード辞書に登録することは、認識語彙から削除することと同等の役割を果たす。

【0090】以上のようにして、キーワード辞書照合に基づいて、発声単位の発声確率を求めることができるのである。

【0091】以上述べた「解析尤度」、「読み尤度」、「単語出現順序」、「モーラ数」、「単語出現頻度」、「キーワード辞書照合」の6種類の項目の何れ一つを用いて求めた発声確率を組み合わせ、発声単位の発声確率とすることができる。一例として、式(6)のような計算式によって発声単位の発声確率を求めることができる。すなわち、
解析尤度を用いて求めた発声確率を P_1
読み尤度を用いて求めた発声確率を P_2
単語出現順序を用いて求めた発声確率を P_3
モーラ数を用いて求めた発声確率を P_4
単語出現頻度を用いて求めた発声確率を P_5
キーワード辞書照合を用いて求めた発声確率を P_6 とすると、

$$P = \sum_{i=1}^6 m_i \cdot P_i \quad \dots(6)$$

ここで、 m_i ：重み係数

【0092】以上のようにして、各発声単位毎に算出された発声確率WSは、図9に示すように、語彙作成部15によって上記第1実施の形態と同様にして作成された認識語彙(発声単位)に付与されて、語彙記憶部17に登録されるのである。

【0093】以上のように、本実施の形態においては、テキスト解析部12は、複数の分割候補が存在する場合には総ての分割候補に解析尤度KSを付与して出力する。また、読み付与部13は、複数の読み方が存在する場合には総ての読み方に読み尤度YSを付与して出力する。さらに、発声確率計算部16によって、上記解析尤度KSおよび読み尤度YS、単語出現順序、モーラ数、単語出現頻度、キーワード辞書照合の少なくとも一つを用いて、得られる発声単位の発声確率を算出する。そして、得られた発声確率WSを認識語彙に付与して語彙記

憶部17に登録するようにしている。

【0094】したがって、本実施の形態によれば、上記第1実施の形態によって作成された音声認識用辞書に登録されている認識語彙中における誤解析による認識語彙や実際には発声されない認識語彙の発声確率を低くすることができ、高い認識精度を得ることができる音声認識用辞書を作成することができる。

【0095】尚、上記関数 $h(i)$ 、 $m(i)$ 、 $f(i)$ およびキーワード辞書の格納場所は特に限定するものではないが、例えば発声確率計算部16の内部メモリに格納してあるものとする。また、説明を簡単にするために、発声確率計算部16を語彙作成部15とは別ブロックに構成して、語彙作成部15で得られた発声単位に関して発声確率を算出するようにしている。しかしながら、発声確率計算部を語彙作成部と同一ブロックに構成して、図2のフローチャートにおける上記ステップS4以降に、発声確率算出動作を組み込んでも差し支えない。

【0096】＜第3実施の形態＞本実施の形態は、上記第2実施の形態における音声認識用辞書作成装置によって作成された音声認識用辞書を搭載した音声認識装置に関するものである。図11は、図6に示す音声認識用辞書作成装置を搭載した音声認識装置のブロック図である。音声認識用辞書作成装置21を構成するテキスト解析部23、読み付与部24、解析辞書メモリ25、語彙作成部26、発声確率計算部27および語彙記憶部28は、上記第2の実施の形態における上記テキスト解析部12、読み付与部13、解析辞書メモリ14、語彙作成部15、発声確率計算部16および語彙記憶部17と同じ構成を有している。そして、語彙記憶部28には、図9に示すように、発声確率が付与された認識語彙が格納されている。

【0097】一方、音声認識装置22は、音響分析部29、尤度計算部30、音響モデル格納部31および照合部32で構成され、マイクに入力された音声信号を語彙記憶部28に格納された認識語彙情報(音声認識用辞書)を用いて認識する。

【0098】上記音響分析部29は、マイクから入力された音声のアナログ波形をデジタル波形に変換し、20msec～40msec程度の短い時間間隔(フレーム)毎に周波数分析し、スペクトルを表すパラメータのベクトル系列に変換する。周波数分析にはLPC(線形予測分析)メルケプストラム等が用いられる。

【0099】上記尤度計算部30は、音響分析部29からの入力音声のパラメータベクトルを用いて、音響モデル格納部31に格納されたHMM(隠れマルコフモデル)等の音韻毎の音響モデルの尤度を計算する。こうして、各音韻の尤度を求めるのである。照合部32は、求められた各音韻の尤度と語彙記憶部28に登録されている総ての認識語彙との照合を行ない、全認識語彙のスコアを算出する。その場合における各認識語彙のスコアは、当

該認識語彙の音韻系列に尤度計算部30で求められた各音韻の尤度を当て嵌めて得た音響尤度と、当該認識語彙に付与されて語彙記憶部28に記憶されている発声確率である言語尤度とを組み合わせて算出する。例えば、認識語彙Wの音響尤度を $p(w)$ とし言語尤度を $q(w)$ とすると、認識語彙Wのスコア $score(w)$ を、

$$score(w) = \alpha \cdot p(w) + \beta \cdot q(w)$$

但し、 α, β ：定数
 によって求めるのである。

【0100】こうして、上記語彙記憶部28に登録されている総ての認識語彙についてスコアを算出し、上位所定値以上のスコアを呈する認識語彙でなる認識候補を認識結果として出力するのである。

【0101】以上のように、本実施の形態においては、音声認識装置は、上記第2実施の形態の音声認識用辞書作成装置によって作成された音声認識用辞書が格納された語彙記憶部28を有している。そして、照合部32によって、尤度計算部30で求められた各音韻の尤度と語彙記憶部28に登録されている総ての認識語彙の音韻系列との照合を行ない、音韻の尤度に基づく音響尤度と上記発声確率である言語尤度とを組み合わせて、全認識語彙のスコアを算出するようにしている。

【0102】したがって、本実施の形態によれば、上記第1実施の形態の場合と同様に、ホームページの呼び出しやテレビのチャンネル切換を音声によって行う場合に、登録されたホームページタイトルや番組名を完全に発声しなくてもホームページの呼び出しやチャンネルの切換を行うことができる。その際に、上記発声確率(言語尤度)に基づくスコアの高い認識語彙を音声認識結果とすることによって、上記入力文字列情報の誤解析によって得られた認識語彙や実際には発声されない認識語彙に誤認識される確率を低減することができる。したがって、高い認識性能を有する音声認識装置を実現することができるのである。

【0103】尚、上記実施の形態においては、上記音声認識装置22は音声認識用辞書作成装置21を搭載している。しかしながら、この発明の音声認識装置は、少なくとも上記音声認識用辞書作成装置によって作成された語彙記憶部28が搭載されていればよく、テキスト解析部23、読み付与部24、解析辞書メモリ25、語彙作成部26および発声確率計算部27は、音声認識装置22とは独立して設けても差し支えない。

【0104】＜第4実施の形態＞本実施の形態は、上記第2実施の形態における音声認識用辞書作成装置によって作成された音声認識用辞書を搭載した音声認識装置の他の例に関するものであり、上記音声認識用辞書作成装置に入力する文字列情報をコンテンツから抽出するものである。

【0105】図12は、図6に示す音声認識用辞書作成装置を搭載した音声認識装置のブロック図である。音声

認識用辞書作成装置41を構成するテキスト解析部43、読み付与部44、解析辞書メモリ45、語彙作成部46、発声確率計算部47および語彙記憶部48、音声認識装置42を構成する音響分析部49、尤度計算部50、音響モデル格納部51および照合部52は、上記第3の実施の形態におけるテキスト解析部23、読み付与部24、解析辞書メモリ25、語彙作成部26、発声確率計算部27、語彙記憶部28、音響分析部29、尤度計算部30、音響モデル格納部31および照合部32と同じ構成を有している。そして、語彙記憶部48には、図9に示すように、発声確率が付与された認識語彙が格納されている。

【0106】取り込み部53は、外部からの文字列情報を含むコンテンツを取り込む。上記コンテンツは、放送によって送信されてくる情報を受信機で受信して取り込んでもよいし、インターネットで配信される情報を通信ネットワーク経由で取り込んでもよいし、光磁気ディスク、磁気テープ、ハードディスクおよびIC(集積回路)カード等の固定的にデータを記録する媒体から取り込んでもよい。文字列情報抽出部54は、取り込み部53で取り込まれたコンテンツの中から音声認識用辞書の作成に使用する文字列情報を抽出する。そして、抽出された文字列情報は、音声認識用辞書作成装置41のテキスト解析部43に送出される。

【0107】抽出条件格納部55には、上記コンテンツに含まれるタグ情報が如何なるタグ情報である場合に文字列情報を抽出するかという抽出条件が格納されている。そして、文字列情報抽出部54は、抽出条件格納部55に格納されている上記抽出条件を参照して、上記音声認識用辞書の作成に用いる文字列情報を抽出するのである。

【0108】図13は、上記取り込み部53および文字列情報抽出部54によって行われる文字列情報抽出処理動作のフローチャートである。以下、図13に従って、文字列情報抽出動作について説明する。ステップS21で、上記取り込み部53によって、上記コンテンツが取り込まれる。以下、文字列情報抽出部54による処理に移行する。ステップS22で、上記取り込まれたコンテンツの先頭文字が読み出される。ステップS23で、上記読み出された文字は空(つまりコンテンツは最後)であるか否かが判別される。その結果、最後であれば文字列情報抽出処理動作を終了する。一方、最後でなければステップS24に進む。ステップS24で、受け取った文字がタグ情報であるか否かが判別される。その結果、タグ情報でなければステップS25に進む。一方、タグ情報であればステップS26に進む。ステップS25で、上記コンテンツにおける次の文字が読み出された後、ステップS23に戻って次の文字の処理に移行する。ステップS26で、抽出条件格納部55の抽出条件が参照されて、上記抽出条件を満たすか否かが判別される。その結果、満たす場合に

はステップS27に進む一方、満たさない場合には上記ステップS25に戻って次の文字の処理に移行する。ステップS27で、上記抽出条件を満たす文字列が抽出されて、テキスト解析部43に送出される。そうした後、上記ステップS25に戻って次の文字の処理に移行する。以下、ステップS23～ステップS27の処理を繰り返し、上記ステップS23においてコンテンツは最後であると判別されと文字列情報抽出処理動作を終了するのである。

【0109】例えば、上記コンテンツがHTML(ハイパーテキスト・マークアップ言語)ファイルであり、抽出条件格納部55に抽出条件「<title>タグが存在する場合は<title>と</title>とで囲まれた文字列を抽出する」が格納されているとすると、入力されたコンテンツ(HTMLファイル)中に、<title>首相官邸のトップページ</title>なる記述があると、文字列「首相官邸のトップページ」が抽出されるのである。

【0110】こうして、文字列情報を含むコンテンツから上記文字列情報抽出部54によって自動的に抽出された文字列に対して、音声認識用辞書作成装置41によって図2に例示するような辞書作成処理を行うことによって、語彙記憶部48に、音声によるホームページ呼び出しに最適な音声認識用辞書が作成・登録されるのである。したがって、音声認識装置42に対して、首相官邸のホームページのタイトル「首相官邸のトップページ」を完全に発声しても、部分的に「首相官邸」や「官邸のトップページ」と発声しても正しく認識され、首相官邸のホームページを呼び出すことができるのである。

【0111】その際に、上記音声認識用辞書作成装置41には、発声確率計算部47が搭載されており、語彙記憶部48には発声確率が付与された認識語彙が登録されている。したがって、文字列情報抽出部54で抽出された文字列に対する誤解析の結果生成された認識語彙や現実には発声されないような認識語彙の発声確率を低くして、高い認識精度を得ることができるのである。

【0112】<第5実施の形態>本実施の形態は、上記第4実施の形態に示す音声認識装置において、音声認識用辞書作成装置に inputs する文字列情報が抽出されるコンテンツを、ウェブページ情報に限定したものである。

【0113】図14は、図6に示す音声認識用辞書作成装置を搭載した音声認識装置のブロック図である。音声認識用辞書作成装置61を構成するテキスト解析部63、読み付与部64、解析辞書メモリ65、語彙作成部66、発声確率計算部67および語彙記憶部68、音声認識装置62を構成する音響分析部69、尤度計算部70、音響モデル格納部71および照合部72は、上記第3の実施の形態におけるテキスト解析部23、読み付与部24、解析辞書メモリ25、語彙作成部26、発声確率計算部27、語彙記憶部28、音響分析部29、尤度計算部30、音響モデル格納部31および照合部32と同じ構成

を有している。そして、語彙記憶部68には、図9に示すように、発声確率が付与された認識語彙が格納されている。

【0114】また、文字列情報抽出部74および抽出条件格納部75は、上記第4実施の形態における文字列情報抽出部54および抽出条件格納部55と同じ構成を有している。そして、文字列情報抽出部74によって抽出された文字列情報は、音声認識用辞書作成装置61のテキスト解析部63に出力される。

【0115】ウェブページ情報取り込み部73は、上記コンテンツとしてのウェブページ情報を取り込んで、先頭文字から順次文字列情報抽出部74に送出する。以後、文字列情報抽出部74によって、抽出条件格納部75に格納されている抽出条件に適合する文字列情報が抽出される。

【0116】一方、制御部76は、上記音声認識装置62による音声認識結果に基づいて、ウェブページの表示制御を行う。そして、ウェブページ表示部77は、制御部76の指示に従ってウェブページを表示するのである。

【0117】図15は、上記ウェブページ情報取り込み部73によって取り込まれるウェブページ情報の一例を示す。ウェブページの情報は、HTML等の言語で記述されている。そして、抽出条件格納部75に抽出条件「<title>タグが存在する場合は<title>と</title>とで囲まれた文字列を抽出する」が格納されているとすると、文字列情報抽出部74は、基本的には図13のフローチャートと同じ文字列情報の抽出動作を行う。すなわち、ウェブページ情報に記述されている文字列を先頭から見ていき、<title>タグを見つけると、<title>と</title>とで挟まれた文字列「鈴木一郎のホームページ」を抽出するのである。

【0118】こうして、ウェブページ情報から上記文字列情報抽出部74によって自動的に抽出された文字列に対して、音声認識用辞書作成装置61によって図2に例示するような辞書作成処理を行うことによって、語彙記憶部68に、音声によるホームページ呼び出しに最適な発声確率付き認識語彙が登録された音声認識用辞書が作成されるのである。その際に、上記文字列「鈴木一郎のホームページ」から作成された発声単位「鈴木」、「一郎」、「鈴木一郎」、「鈴木一郎のホームページ」、「一郎のホームページ」には、図16に示すように当該ホームページのURL(<http://www.suzuki.xxx.jp>等)が付与されて語彙記憶部68に登録するようにしている。

【0119】したがって、上記音声認識装置62に対して「鈴木一郎」と発声された場合は、音響分析部69で入力音声が一ベクトル系列に変換され、尤度計算部70で上記ベクトル系列に基づいて各音韻の尤度が算出され、照合部72で語彙記憶部68の語彙との照合を行って「鈴木一郎」という語彙として認識されると共に、認識語彙

「鈴木一郎」に付与されているURL(<http://www.suzuki.xxx.jp>)が得られるのである。

【0120】そうすると、上記制御部76は、得られたURLに基づいて「鈴木一郎のホームページ」にアクセスして「鈴木一郎のホームページ」のウェブページ情報を取得し、ウェブページ表示部77に当該ウェブページの表示を指示して表示させるのである。

【0121】すなわち、本実施の形態によれば、ウェブページの情報から自動的に発声確率付きの発声単位を生成して適切な音声認識用辞書を作成することができる。つまり、図15のようなウェブページから図16に示すような音声認識用辞書を自動的に作成することができる。したがって、ウェブページのタイトル「鈴木一郎のホームページ」を完全に発声しても、部分的に「鈴木」や「鈴木一郎」や「一郎のホームページ」と発声しても正しく認識し、「鈴木一郎のホームページ」のウェブページ情報を取得してウェブページ表示部77に表示することができるのである。

【0122】上記ブラウザのブックマークやお気に入り等でホームページをブラウザに登録した場合、利用者が見て判断する情報としてタイトルを利用する場合がある。その場合のタイトルは、飽くまでも見るための情報であって発声するための文字列情報ではないため、極端に長いタイトルが与えられる場合がある。そのような場合でも、<title>タグで囲まれた文字列を抽出して上述した方法で音声認識用辞書を作成することによって、より短い発声で上記ブックマークやお気に入りに登録したホームページを呼び出すことが可能になるのである。

【0123】尚、本実施の形態においては、上記<title>タグを用いる場合を例に説明したが、<title>タグ以外にも文字フォントを変更するタグ等、用途に応じた抽出条件を抽出条件格納部75に格納しておくができる。そうすることによって、様々な用途に応じた音声認識用辞書を作成することが可能になる。

【0124】また、上記抽出条件格納部75に、上記発声確率の変更条件を格納することも可能である。その場合の発声確率の変更条件は、例えば、「URLの表記を基にタイトルから発声確率の高い発声単位を選ぶ」である。具体的には、「URL中に含まれる読みに関するアルファベット表記と、テキスト解析部63および読み付与部64によるテキスト解析および読み付与の結果から得られた読みの表記(アルファベット表記)の一致具合を調べ、一致する単語があればその単語(読み方)を含む発声単位の発声確率を高くする」のである。例えば、タイトル「朝日新聞のホームページ」とそのURL「<http://www.asahi.com>」とを例に取ると、URL中に含まれる読みに関するアルファベット表記「asahi」とテキスト解析および読み付与の結果得られた単語「朝日」の読みの表記「asahi」とが同一である。したがって、単語「朝日」を含む発声単位の発声確率を高く設定することができる。こ

して、タイトル「朝日新聞のホームページ」に関する高い認識精度を可能にする音声認識用辞書を簡単に作成することが可能になるのである。

【0125】<第6実施の形態>本実施の形態は、上記第4実施の形態に示す音声認識装置において、音声認識用辞書作成装置に入力する文字列情報が抽出されるコンテンツを、テレビ番組情報に限定したものである。

【0126】図17は、図6に示す音声認識用辞書作成装置を搭載した音声認識装置のブロック図である。音声認識用辞書作成装置81を構成するテキスト解析部83、読み付与部84、解析辞書メモリ85、語彙作成部86、発声確率計算部87および語彙記憶部88、音声認識装置82を構成する音響分析部89、尤度計算部90、音響モデル格納部91および照合部92は、上記第3の実施の形態におけるテキスト解析部23、読み付与部24、解析辞書メモリ25、語彙作成部26、発声確率計算部27、語彙記憶部28、音響分析部29、尤度計算部30、音響モデル格納部31および照合部32と同じ構成を有している。そして、語彙記憶部88には、図9に示すように、発声確率が付与された認識語彙が格納されている。

【0127】また、文字列情報抽出部94および抽出条件格納部95は、上記第4実施の形態における文字列情報抽出部54および抽出条件格納部55と同じ構成を有している。そして、文字列情報抽出部94によって抽出された文字列情報は、音声認識用辞書作成装置81のテキスト解析部83に出力される。

【0128】テレビ番組情報取り込み部93は、上記コンテンツとしてのテレビ番組情報を取り込んで、先頭文字から順次文字列情報抽出部94に送出する。上記テレビ番組情報の取り込みは、1日分あるいは1週間分等のまとまった電子番組データを取り込むことによって行う。尚、この電子番組データの取り込みは、文字放送受信機で受信して取り込んでもよいし、インターネット等のネットワーク経由で取り込んでもよい。また、光磁気ディスク等の記録メディアから取り込んでも差し支えない。そうした後、文字列情報抽出部94によって、抽出条件格納部95に格納されている抽出条件に適合する文字列情報が抽出される。

【0129】一方、制御部96は、上記音声認識装置82による音声認識結果に基づいて、テレビの表示や録画・再生を制御する。そして、テレビ表示部97は、制御部96の指示に従ってテレビの映像を表示する。録画部98は、制御部96の指示に従ってテレビ番組を録画する。再生部99は、制御部96の指示に従って録画部98で録画されたテレビ番組を再生する。

【0130】テレビ番組表には、番組の日時、チャンネル、番組名等の情報が、決められた形式に従って記載されている。したがって、上記5実施の形態と同様に、テレビ番組表の中から番組名等の特定項目の文字列

を抽出することによって、音声認識用辞書を作成することができる。

【0131】図18は、上記テレビ番組情報取り込み部93によって取り込まれるテレビ番組情報の一例を示す。抽出条件格納部95には、抽出条件「番組名というタグが付いている文字列を抽出する」が格納されているものとする。そうすると、文字列情報抽出部94は、基本的には図13のフローチャートと同じ文字列情報の抽出動作を行う。すなわち、テレビ番組情報に記述されている文字列を先頭から見ていき、タグ「番組名」を見つければ、そのタグ「番組名」に該当する文字列「NHKニュースおはよう日本」を抽出するのである。

【0132】こうして、テレビ番組情報から上記文字列情報抽出部94によって自動的に抽出された文字列に対して、音声認識用辞書作成装置81によって図2に例示するような辞書作成処理を行うことによって、語彙記憶部88に、音声によるテレビ番組表示に最適な発声確率付き認識語彙が登録された音声認識用辞書が作成されるのである。その際に、上記文字列「NHKニュースおはよう日本」から作成された発声単位「NHK」、「ニュース」、「NHKニュース」、「おはよう日本」、「NHKニュースおはよう日本」には、図19に示すように当該番組のチャンネル情報（「NHK総合」等）、日付情報および時刻情報等が付与されて語彙記憶部88に登録するようにしている。

【0133】したがって、上記音声認識装置82に対して「NHKニュース」と発声された場合は、音響分析部89で入力音声が入力ベクトル系列に変換され、尤度計算部90で上記ベクトル系列に基づいて各音韻の尤度が算出され、照合部92で語彙記憶部88の語彙との照合を行って「NHKニュース」という語彙として認識されると共に、認識語彙「NHKニュース」に付与されているチャンネル情報（NHK総合）、日付情報（5月5日）および時刻情報（5:00～8:15）が得られるのである。

【0134】そうすると、上記制御部96は、得られたチャンネル情報、日付情報および時刻情報に基づいて、テレビ表示部97にテレビの表示を指示して、5月5日の午前5時に「NHKニュースおはよう日本」を表示させるのである。

【0135】すなわち、本実施の形態によれば、テレビ番組情報から自動的に発声確率付きの発声単位を生成して適切な音声認識用辞書を作成することができる。つまり、図18に示すようなテレビ番組情報から図19に示すような音声認識用辞書を自動的に作成することができる。したがって、テレビ番組の番組名「NHKニュースおはよう日本」を完全に発声しても、部分的に「NHK」や「おはよう日本」や「NHKニュースおはよう」と発声しても正しく認識し、テレビ番組「NHKニュースおはよう日本」のチャンネル情報、日付情報および時刻情報を取得してテレビ表示部97に表示することができるのである。

る。

【0136】尚、本実施の形態においては、上記番組名タグを用いる場合を例に上げて説明したが、用途に応じて様々な抽出条件を抽出条件格納部95に格納しておくができる。例えば、録画を予約する場合は、現在の日付及び時間よりも先の番組が認識対象になるはずであるから、抽出条件格納部95に「日付が本日以降（または、時間が現在時刻以降）の番組名を抽出する」等の抽出条件を格納しておくことも可能である。

【0137】＜第7実施の形態＞本実施の形態は、上記第6実施の形態の利用例に関し、音声リモコンによって情報家電機器を制御するシステムに関するものである。この情報家電機器音声リモコン制御システムは、図20に示すような構成を有している。音声リモコン101は、マイク102、スピーカー103およびリモコン制御部104で構成される。そして、マイク102に音声が入力されると、リモコン制御部104によって、入力音声は通信回線105を経由して音声認識装置106に送出される。スピーカー103は、通信回線105を経由して音声認識装置106から送出されてくる認識結果等を音声で出力し、認識結果の確認等に用いられる。尚、通信回線105には、例えば赤外線等を用いる。

【0138】音声認識装置106は、図17に示す音声認識用辞書作成装置を搭載した音声認識装置の構成から制御部96、テレビ表示部97、録画部98および再生部99を削除した構成を有している。そして、通信回線105から音響分析部（図示せず）に入力された音声を認識し、認識結果を通信回線107を経由して情報家電機器108に送出する。その場合に、音声認識用辞書作成装置のテレビ番組情報取り込み部（図示せず）に入力されるテレビ番組情報は、通信回線107を経由して情報家電機器108から入力される。尚、通信回線107は、上述のごとく音声認識装置106と情報家電機器108とを繋ぐネットワークであり、有線や無線のLAN（ローカルエリアネットワーク）等のネットワークである。具体的には、音声認識装置106をパーソナルコンピュータ、情報家電機器108をテレビと考えることができる。尚、本実施例においては、音声認識装置106と情報家電機器108とを通信回線107で繋ぐ構成にしているが、音声認識装置106を情報家電機器108内に組み込んでも一向に構わない。

【0139】上記情報家電機器108は、上記通信回線107からの情報に基づいて各種制御を行う機器制御部109と、映像や音楽等の当該機器の出力情報となるメイン情報を記憶するメイン情報メモリ110と、映像の番組名や音楽のタイトル名等の上記メイン情報に関わるサブ情報を記憶するサブ情報メモリ111と、ディスプレイやスピーカー等を有して上記メイン情報を出力するメイン情報出力部112を搭載している。

【0140】通信回線113は、上記情報家電機器10

8と屋外の情報提供センター114とを繋ぐネットワークであり、電話回線やケーブルテレビ回線等のネットワークおよびデジタル放送網である。情報提供センター114は、プロバイダーや放送局に相当するものであり、情報家電機器108に対して通信回線113を介してメイン情報を提供する。この情報提供センター114は、情報家電機器108に送出すべきメイン情報を記憶するメイン情報メモリ115と、メイン情報に関わるサブ情報を記憶するサブ情報メモリ116と、各メモリ115, 116に対する各種情報の記憶や読み出しおよび読み出し情報の通信回線113への送出等の制御を行う制御部117を有している。

【0141】上記構成を有する情報家電機器音声リモコン制御システムにおいては、音声リモコン(ワイヤレスマイク)101に入力された音声によってメイン情報を選択して、最寄の情報家電機器108で再生することができる。以下、家庭において音声リモコン101に向かって、例えば「NHKニュース」と発声してテレビ画面に「NHKニュース」を出力させる場合を例に、本情報家電機器音声リモコン制御システムの動作を、具体的に説明する。

【0142】上記情報提供センター114が提供するメイン情報としてのテレビ映像情報およびサブ情報としてのテレビ番組情報は、通信回線113を経由して情報家電機器108に取り込まれ、機器制御部109によってテレビ番組情報がサブ情報メモリ111に格納される。そして、機器制御部109は、サブ情報メモリ111の記憶内容(テレビ番組情報)を通信回線107を経由して音声認識装置106に送出する。

【0143】上記音声認識装置106は、受け取ったテレビ番組情報を上記音声認識用辞書作成装置のテレビ番組情報取り込み部に取り込んで、図17の説明で述べたようにして文字列情報の抽出と音声認識用辞書の作成を行う。その結果、作成された音声認識用辞書には、図19に示すようなテレビ番組指定用の認識語彙が発声確率とチャンネル情報、日付情報、時刻情報とが付与されて登録される。

【0144】この状態で、上記音声リモコン101のマイク102に対して「NHKニュース」と発声されると、入力された音声は通信回線105を経由して音声認識装置106の音響分析部に取り込まれ、上述のようにして、ベクトル系列への変換、各音韻の尤度算出および上記音声認識用辞書との照合が行われ、「NHKニュース」という語彙として認識されると共に、認識語彙「NHKニュース」に付与されているチャンネル情報、日付情報および時刻情報が得られる。そして、これらの情報をコマンドとして通信回線107を経由して情報家電機器108に送信するのである。

【0145】そうすると、上記情報家電機器108の機器制御部109は、受信したコマンドをサブ情報メモリ

111の内容と照合して解釈し、上記日付情報および時刻情報による日付および時刻になると、上記チャンネル情報によるNHK総合の映像情報がメイン情報出力部112に出力されるのである。尚、受信したコマンドが録画コマンドである場合には、上記映像情報がメイン情報メモリ110に格納される。また、受信したコマンドが再生コマンドである場合には、メイン情報メモリ110に格納された上記映像情報が読み出されて、メイン情報出力部112に出力される。

【0146】このように、本実施の形態によれば、上記音声認識装置106の音響分析部に対する音声入力を、音声リモコン101から赤外線等の通信回線105を経由して行うようにしている。したがって、情報家電機器音声制御システムの操作性を多に向上できるのである。

【0147】尚、図20に示す情報家電機器音声リモコン制御システムは、上記音声認識装置106は図17に示す音声認識装置であって、サブ情報としてテレビ番組情報が入力される場合を例に説明している。しかしながら、この発明はこれに限定されるものではなく、音声認識装置106を図14に示す音声認識装置とし、サブ情報としてウェブページ情報を用いてもよい。あるいは、音声認識装置106を図12に示す音声認識装置とし、サブ情報として一般のコンテンツ情報を用いてもよい。

【0148】＜第8実施の形態＞上記各実施の形態によれば、テキスト解析の結果得られる単語の組合せが異なる複数の接続単語で成る発声単位は、各発声単位間の音響的な類似度には何ら考慮することなく、認識語彙としてそのまま音声認識用辞書に登録している。したがって、ホームページのタイトルに「首相官邸のトップページ」と「司法省の主張」とがある場合に、発声「首相…」が「主張…」と誤認識されて「司法省の主張」のホームページが表示されてしまう場合がある。本実施の形態は、このような場合に対処するものである。

【0149】図21は、図6に示す音声認識用辞書作成装置を搭載した音声認識装置のブロック図である。音声認識用辞書作成装置121を構成するテキスト解析部123、読み付与部124、解析辞書メモリ125、発声確率計算部127および語彙記憶部128、音声認識装置122を構成する音響分析部129、尤度計算部130、音響モデル格納部131および照合部132は、上記第3の実施の形態におけるテキスト解析部23、読み付与部24、解析辞書メモリ25、発声確率計算部27、語彙記憶部28、音響分析部29、尤度計算部30、音響モデル格納部31および照合部32と同じ構成を有している。そして、語彙記憶部128には、図9に示すように、発声確率が付与された認識語彙が格納されている。

【0150】類似度計算部133は、語彙作成部126によって上記第3の実施の形態における語彙作成部26と同様にして作成された任意の2つの発声単位の音響的

な類似度を計算する。その結果、例えば発声単位「首相(しゅしょう)」と発声単位「主張(しゅちょう)」との方が、発声単位「首相(しゅしょう)」と発声単位「官邸(かんでい)」よりも音響的に類似していることが分る。尚、上記類似度の計算には様々な方法があるが、例えば一致する音素が幾つ存在するかに基づいて計算することで実現できる。

【0151】そして、語彙作成部126は、上記発声確率計算部127で計算された各発声単位の発声確率値を、類似計算部133による類似度の計算結果を用いて変更するのである。

【0152】ここで、ホームページのタイトルに「首相官邸のトップページ」と「司法省の主張」とがあるとす。テキスト解析部123から語彙作成部126までによる一連の処理によって、発声単位「首相(しゅしょう)」、「官邸(かんでい)」、「司法省(しほうしょう)」、「主張(しゅちょう)」等が発声確率と共に求められる。その場合、発声単位「首相」と発声単位「主張」との発声確率を比べると、前者の方が高い値が付与されることは、上記第2実施の形態で述べた通りである。しかしながら、発声単位「主張」もそれなりの発声確率を有しているため、そのまま語彙記憶部128に登録して音声認識用辞書とすると、発声「首相」が「主張」と誤認識されて「司法省の主張」のホームページが表示されてしまう恐れが多分にある。

【0153】そこで、上記語彙作成部126は、類似度計算部133による類似度に基づいて、音響的に類似した2つの発声単位を検索する。そして、一方の発声単位がホームページタイトルにおいて中心的な役割を果たし、他方の発声単位がそうでない場合には、中心的な発声単位の発声確率を更に上げ、そうでない方の発声単位の発声確率を更に下げるのである。その場合、該当発声単位が中心的な役割を果たすか否かは、発声確率計算部127で与えられた確率値の大小で判断することができる。尚、その際に、発声確率値を如何様に変更するかは、特に規定するものではない。例えば、発声確率値の値に応じて変更量を設定してもよいし、類似度に基づいて変更量を設定してもよい。また、中心的な役割を果たしていない発声単位の発声確率値を「0」に変更することも可能である。

【0154】以上のように、本実施の形態においては、語彙作成部126によって生成された任意の2つの発声単位の類似度を類似度計算部133で計算し、音響的に類似した2つの発声単位が存在する場合には、語彙作成部126は、音響的に類似した2つの発声単位のうち中心的な役割でない発声単位の確率を低くするようにしている。したがって、本実施の形態によれば、上記中心的な役割を果たす発声単位を発声した際に、上記中心的な役割を果たしていない発声単位に誤認識されることを防止することができる。すなわち、本実施の形態によれば、

さらに認識性能の高い音声認識装置を構築することができるのである。

【0155】<第9実施の形態>上記各実施の形態においては、語彙記憶部6,17,28,48,68,88,128には、住所録やウェブページ情報やテレビ番組情報から生成された発声単位が認識用の語彙として登録されている。したがって、語彙記憶部6,17,28,48,68,88,128は、登録されている住所やホームページのタイトルや番組名が発声された場合には高い認識率を示す。ところが、ユーザは、常に語彙記憶部6,17,28,48,68,88,128に登録されている住所やタイトル名や番組名を発声するとは限らない。したがって、登録されていない住所やタイトル名や番組名が発声された場合の認識精度が、極端に低下するという問題が生ずることになる。本実施の形態は、このような場合に対処するものである。

【0156】図22は、本実施の形態における音声認識装置のブロック図である。音声認識装置141を構成する音響分析部142、尤度計算部143、音響モデル格納部144および第1語彙記憶部146は、上記第3の実施の形態における音声認識装置22を構成する音響分析部29、尤度計算部30、音響モデル格納部31および語彙記憶部28と同じ構成を有している。そして、第1語彙記憶部146には、図16に示すように、発声確率が付与されたホームページタイトル用の認識語彙が格納されているものとする。

【0157】第2語彙記憶部147は、上述した各音声認識用辞書作成装置によって特定のテキストに基づいて作成されるのではなく、固定した一般的な語彙が登録された音声認識用辞書であり、第1語彙記憶部146と同様に照合部145によって照合される。つまり、第2語彙記憶部147は、第1語彙記憶部146は利用者が選択した語彙だけを登録しているのに対して、固定された一般的な語彙が登録されている点において異なるのである。

【0158】上記照合部145は、第1語彙記憶部146および第2語彙記憶部147の語彙との照合を行い、第1語彙記憶部146に登録されている語彙が認識語彙であると判断した場合は、その語彙および当該語彙に付与されているサブ情報(URL)を認識結果として出力する。一方、第2語彙記憶部147に登録されている語彙が認識語彙であると判断した場合は、その語彙を認識結果として検索部148に送出する。そうすると、検索部148は、照合部145から受け取った認識結果に対応する文字列情報を、第1語彙記憶部146に登録されている語彙が住所録から生成された発声単位である場合には住所録から、ウェブページ情報から生成された発声単位である場合にはウェブページ情報から、テレビ番組情報から生成された発声単位である場合にはテレビ番組情報から検索する。選択部149は、検索部148が受け

取った語彙(上記認識結果)と上記検索された文字列情報との中から、利用者が第1語彙記憶部146に登録する語彙情報と付与情報とを選択して第1語彙記憶部146に登録するためのものである。

【0159】以下、上記第1語彙記憶部146に登録されている語彙はウェブページ情報から生成された発声単位である場合を例に、本実施の形態における音声認識装置の動作について説明する。その場合、本音声認識装置においては、第1語彙記憶部146に登録されている語彙を発声することで、認識結果に応じたウェブページを表示することができる。第1語彙記憶部146には、少なくとも認識語彙/読み方のペアと当該認識語彙に関する文字列情報(URL)とが対応付けられて登録されている。その場合、登録されている認識語彙は、上記第5実施の形態で述べた方法によって作成されたものでもよいし、利用者が自分で入力したものでもよい。何れしても、利用者がよく見るウェブページに限定して、それに関する認識語彙を登録したものが第1語彙記憶部146である。

【0160】ところで、利用者が発声する語彙は常に第1語彙記憶部146にあるものだけであるとは限らない。そこで第1語彙記憶部146には無い語彙を発声した場合でも認識結果が得られるように、第2語彙記憶部147が設けられている。この第2語彙記憶部147には、利用者が登録した認識語彙だけではなく、一般に使用される任意の語彙が記憶されている。第1語彙記憶部146の登録内容が可変で一般にある範囲に限定された少数であるのに対して、第2語彙記憶部147の登録内容は固定で一般にある程度広い範囲にわたって大量である。

【0161】上記照合部145は、認識時において第1語彙記憶部146の登録内容と照合するか第2語彙記憶部157の登録内容と照合するかを判断する。その方法に付いては特に限定するものではなく、例えば、先ず第1語彙記憶部146と照合を行い、認識尤度が十分高くない場合に第2語彙記憶部147との照合を行う方法でもよいし、第1語彙記憶部146と第2語彙記憶部147との両方で照合を行い、認識尤度の高い認識候補を認識結果とする方法でもよい。

【0162】ここで、上記第1語彙記憶部146に図16に示すような認識語彙/読み/発声確率/URLの組が登録されている場合に、使用者が「佐藤」と発声したとする。その場合、第1記憶部146には読み「satou」は存在しないので、発声「佐藤」は認識されない。その場合には、第2語彙記憶部147との照合で「佐藤」という語彙が認識されることになる。こうして得られた認識結果「佐藤」は、検索部148に送出される。そして、検索部148によって、キーワード「佐藤」でウェブページの検索が行われる。尚、このウェブページの検索は、インターネットで普及している検索エンジンを用いることで実

現できる。検索エンジンは、与えられたキーワードから、それに関連するウェブページのURLを探し出すプログラムである。一般に、探し出されるURLは複数個あり、検索エンジンはそれらを利用者に提示する。そして、選択部149は、利用者によって、検索部148による検索結果の中から所望の「佐藤さんのホームページ」のURLが選択されると、認識語彙「佐藤/satou」と「佐藤さんのホームページ」のURLとを関連付けて第1語彙記憶部146に登録するのである。

【0163】こうして、上記第1語彙記憶部146に認識語彙「佐藤/satou」と「佐藤さんのホームページ」のURLとが登録されると、以降は検索や選択を行うことなく、「佐藤」と発声すれば直ちに所望のウェブページ「佐藤さんのホームページ」を見ることができるようになるのである。また、第1語彙記憶部146に対する新たな認識語彙の登録も、文字を入力したりすることなく、発声と選択部149による選択だけで簡単に行うことができるのである。

【0164】上記各実施の形態における音声認識用辞書作成装置で作成された音声認識用辞書を用いる音声認識装置は、携帯電話や電子手帳等の携帯端末器に搭載することが有効である。すなわち、このような携帯端末器においては、操作指示を行う場合にはキー操作よりも発声による方が操作性がよい。ところが、出先等においては操作指示を行うための文言を予め決められている通りに正確に発声するのは困難であり、そのような場合に対処するための音声認識用辞書を利用者が作成するのは更に困難である。

【0165】上記各実施の形態における音声認識用辞書作成装置によれば、一つの入力文字列から、総ての分割候補、総ての読み候補、総ての連接単語の組み合わせを考慮して、複数の発声単位でなる認識用語彙を自動的に生成できるため、予め設定された文字列の部分文字列を発声しても正しく認識できる音声認識用辞書を非常に簡単に作成することができる。したがって、このような音声認識用辞書を用いる音声認識装置を搭載することは、携帯端末器用の音声合成装置として非常に有効なのである。

【0166】ところで、上記各実施の形態における上記テキスト解析部、読み付与部、語彙作成部、語彙記憶部、発声確率計算部、取り込み部、文字列情報抽出部および類似度計算部による上記解析手段、読み付与手段、語彙作成手段、語彙記憶手段、発声確率算出手段、取り込み手段、文字列情報抽出手段及び類似度算出手段としての機能は、プログラム記録媒体に記録された辞書作成処理プログラムによって実現される。上記実施の形態における上記プログラム記録媒体は、ROM(リード・オンリ・メモリ)でなるプログラムメディアである。あるいは、外部補助記憶装置に装着されて読み出されるプログラムメディアであってもよい。尚、何れの場合においても、上記プログラ

ムメディアから辞書作成処理プログラムを読み出すプログラム読み出し手段は、上記プログラムメディアに直接アクセスして読み出す構成を有していてもよいし、RAM(ランダム・アクセス・メモリ)に設けられたプログラム記憶エリア(図示せず)にダウンロードし、上記プログラム記憶エリアにアクセスして読み出す構成を有していてもよい。尚、上記プログラムメディアからRAMの上記プログラム記憶エリアにダウンロードするためのダウンロードプログラムは、予め本体装置に格納されているものとする。

【0167】ここで、上記プログラムメディアとは、本体側と分離可能に構成され、磁気テープやカセットテープ等のテープ系、フロッピー(登録商標)ディスク、ハードディスク等の磁気ディスクやCD(コンパクトディスク)・ROM、MO(光磁気)ディスク、MD(ミニディスク)、DVD(デジタルビデオディスク)等の光ディスクのディスク系、IC(集積回路)カードや光カード等のカード系、マスクROM、EPROM(紫外線消去型ROM)、EEPROM(電氣的消去型ROM)、フラッシュROM等の半導体メモリ系を含めた、固定的にプログラムを担持する媒体である。

【0168】また、上記各実施の形態における音声認識装置は、モデムを備えてインターネットを含む通信ネットワークと接続可能な構成を有していれば、上記プログラムメディアは、通信ネットワークからのダウンロード等によって流動的にプログラムを担持する媒体であっても差し支えない。尚、その場合における上記通信ネットワークからダウンロードするためのダウンロードプログラムは、予め本体装置に格納されているものとする。あるいは、別の記録媒体からインストールされるものとする。

【0169】尚、上記記録媒体に記録されるものはプログラムのみに限定されるのではなく、データも記録することが可能である。

【0170】

【発明の効果】以上より明らかなように、第1の発明の音声認識用辞書作成装置は、一つの文字列情報から解析手段によって得られた総ての分割候補および読み付与手段によって得られた総ての読み候補に基づいて、語彙作成手段によって、読みが付与された1つまたは複数の発声単位を生成し、生成された発声単位を認識語彙として語彙記憶手段によって登録するので、与えられた文字列情報から、発声の可能性のある発声単位を認識語彙とする音声認識用辞書を生成できる。したがって、利用者が、予め設定された文字列中のどの部分文字列を発声しても正しく認識するための音声認識用辞書を、低コストで作成することができる。

【0171】また、上記第1の発明の音声認識用辞書作成装置は、発声確率算出手段によって、解析尤度、読み尤度、単語出現順序、モーラ数、単語出現頻度及びキーワード

ード辞書照合結果の少なくとも一つを用いて上記生成された各発声単位の発声確率を計算し、上記語彙作成手段によって、上記各発声単位でなる認識語彙を、上記発声確率を付与して上記語彙記憶手段に記憶させれば、上記音声認識用辞書中に、上記解析手段による誤解析の結果生成された認識語彙や実際には発声されない認識語彙が登録されていても、このような不要な認識語彙の発声確率を小さく設定して、高い認識精度を実現できる音声認識用辞書を作成できる。

【0172】また、上記第1の発明の音声認識用辞書作成装置は、文字列情報を含むコンテンツを取り込む取り込み手段と、辞書作成に必要な文字列情報の抽出条件が格納された抽出条件格納手段と、上記抽出条件を参照して上記コンテンツから文字列情報を抽出して上記解析手段に送出する文字列情報抽出手段を備えれば、上記コンテンツ情報から自動的に上記音声認識用辞書を作成することができる。

【0173】また、上記第1の発明の音声認識用辞書作成装置は、上記取り込み手段を、上記コンテンツとしてウェブページの情報を取り込むように成せば、上記抽出条件格納手段に、例えば、「<title>タグが存在する場合は<title>と</title>とで囲まれた文字列を抽出する」を格納しておくことによって、ウェブページのタイトル「<title>首相官邸のトップページ</title>」から文字列「首相官邸のトップページ」を抽出して上記音声認識用辞書を自動的に作成できる。

【0174】また、上記第1の発明の音声認識用辞書作成装置は、上記取り込み手段を、上記コンテンツとしてテレビ番組の情報を取り込むように成せば、上記抽出条件格納手段に、例えば、「番組名というタグが付いている文字列を抽出する」を格納しておくことによって、タグ「番組名」に該当する文字列「NHKニュースおはよう日本」を抽出して上記音声認識用辞書を自動的に作成できる。

【0175】また、上記第1の発明の音声認識用辞書作成装置は、上記語彙作成手段で生成された各発声単位間の音響的な類似度を計算する類似度算出手段を備え、上記語彙作成手段によって、上記各認識語彙に付与する発声確率を上記類似度に応じて変更するように成せば、発声単位「首相(しゅしょう)」と発声単位「主張(しゅちょう)」とが音響的に類似している場合には、例えば、上記発声確率の値が高くて入力文字列中において中心的な役割を果たしている発声単位「首相」の発声確率の値を更に高める一方、そうでない発声単位「主張」の発声確率の値を更に低めることによって、中心的な役割を果たす発声「首相」が「主張」と誤認識されることを防止できる。

【0176】また、第2の発明の音声認識用辞書作成方法は、入力された文字列情報を解析して構成単語に分割し、総ての分割候補を出力するステップと、上記分割された各構成単語に読みを付与して総ての読み候補を出力

するステップと、上記総ての分割候補および読み候補に基づいて、読みが付与された1つまたは複数の発声単位を認識語彙として生成するステップと、上記生成された各認識語彙を音声認識用辞書として記憶するステップを備えたので、上記第1の発明の場合と同様に、与えられた文字列情報から、発声の可能性がある発声単位を認識語彙とする音声認識用辞書を生成できる。したがって、予め設定された文字列中のどの部分文字列を発声しても正しく認識するための音声認識用辞書を作成することができる。

【0177】また、第3の発明の音声認識装置は、照合用の辞書として、上記第1の発明の音声認識用辞書作成装置によって作成された音声認識用辞書を用いるので、与えられた文字列情報から生成された発声の可能性がある発声単位を認識語彙とする音声認識用辞書を用いることによって、予め設定された文字列の部分文字列を発声しても正しく認識することができる。

【0178】また、上記第3の発明の音声認識装置は、上記ウェブページの情報を取り込む音声認識用辞書作成装置によって作成された音声認識用辞書を上記辞書として用い、制御手段によって、認識結果に基づいてウェブページ表示手段の表示内容を切換え制御して、上記認識結果に応じたウェブページを表示するようにすれば、発声内容に応じたウェブページを正しく表示することができる。

【0179】また、上記第3の発明の音声認識装置は、上記文字情報化されたテレビ番組の情報を取り込む音声認識用辞書作成装置によって作成された音声認識用辞書を上記辞書として用い、制御手段によって、認識結果に基づいてテレビ表示手段、録画手段および再生手段を制御するようにすれば、発声内容に応じて表示チャンネルの切換え、録画条件の設定または録画番組の再生を正しく行うことができる。

【0180】また、上記第3の発明の音声認識装置は、上記第1の発明の音声認識用辞書作成装置によらずに作成された補助辞書を備えて、照合手段によって上記辞書および補助辞書との照合を行うようにすれば、上記辞書に登録されていない語彙を発声した場合でも、その語彙を正しく認識することができる。さらに、上記照合の結果、認識結果として、上記補助辞書の認識語彙が選択された場合には、検索手段によって、上記音声認識用辞書作成装置に入力された例えばウェブページのタイトルに関連するウェブページ情報の中から当該認識結果に該当する文字列を検索し、選択手段によって、上記検索された複数の文字列の中から上記辞書に登録する文字列を選択すれば、その文字列を上記辞書に登録することによって認識語彙を増やすことができ、認識速度を向上できるのである。

【0181】また、第4の発明の音声認識装置は、上記第1の発明の音声認識用辞書作成装置を搭載し、上記音

声認識用辞書作成装置によって作成された音声認識用辞書を上記照合用の辞書として用いるので、上記搭載された音声認識用辞書作成装置に文字列情報を入力することによって、予め設定された文字列中のどの部分文字列を発声しても正しく認識するための音声認識用辞書を自動的に作成できる。したがって、高い認識精度を得ることができる。

【0182】また、第5の発明の携帯端末器は、上記第3、第4の発明の音声認識装置を搭載したので、出先等において、操作指示を行うための文言を予め決められている通りに正確に発声しなくとも、例えばホームページの音声呼び出し等を正しく行うことができる。

【0183】また、第6の発明のプログラム記録媒体は、コンピュータを、上記第1の発明における解析手段、読み付与手段、語彙作成手段および語彙記憶手段として機能させる辞書作成処理プログラムが記録されているので、上記第1の発明の場合と同様に、読みが付与された1つまたは複数の発声単位を生成し、認識語彙として登録することができる。したがって、予め設定された文字列中のどの部分文字列を発声しても正しく認識するための音声認識用辞書を作成することができる。

【図面の簡単な説明】

【図1】 この発明の音声認識用辞書作成装置におけるブロック図である。

【図2】 図1に示す音声認識用辞書作成装置によって実行される辞書作成処理動作のフローチャートである。

【図3】 図1におけるテキスト解析部の出力結果の一例を示す図である。

【図4】 図4における読み付与部による読み付与結果の一例を示す図である。

【図5】 図1に示す語彙記憶部に作成される音声認識用辞書の一例を示す図である。

【図6】 図1とは異なる音声認識用辞書作成装置におけるブロック図である。

【図7】 図6におけるテキスト解析部の出力結果の一例を示す図である。

【図8】 図6における読み付与部の出力結果の一例を示す図である。

【図9】 図6における語彙作成部の出力結果の一例を示す図である。

【図10】 モーラ数 i を変数とする関数 $m(i)$ の概念を示す図である。

【図11】 図6に示す音声認識用辞書作成装置を搭載した音声認識装置のブロック図である。

【図12】 音声認識用辞書作成装置を搭載した図11とは異なる音声認識装置のブロック図である。

【図13】 図12における取り込み部および文字列情報抽出部によって行われる文字列情報抽出処理動作のフローチャートである。

【図14】 音声認識用辞書作成装置を搭載した図11

および図12とは異なる音声認識装置のブロック図である。

【図15】 ウェブページ情報の一例を示す図である。

【図16】 図14における語彙記憶部に作成される音声認識用辞書の一例を示す概念図である。

【図17】 音声認識用辞書作成装置を搭載した図11、図12および図14とは異なる音声認識装置のブロック図である。

【図18】 テレビ番組情報の一例を示す図である。

【図19】 図17における語彙記憶部に作成される音声認識用辞書の一例を示す概念図である。

【図20】 情報家電機器音声リモコン制御システムのブロック図である。

【図21】 音声認識用辞書作成装置を搭載した図11、図12、図14および図17とは異なる音声認識装置のブロック図である。

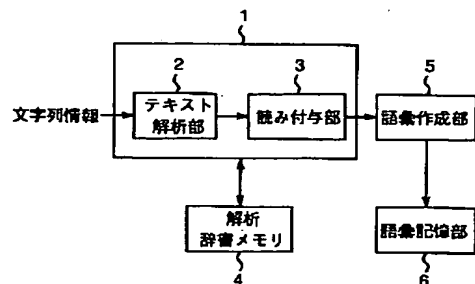
【図22】 音声認識装置のブロック図である。

【符号の説明】

1, 11…解析処理部、
2, 12, 23, 43, 63, 83, 123…テキスト解析部、
3, 13, 24, 44, 64, 84, 124…読み付与部、
4, 14, 25, 45, 65, 85, 125…解析辞書メモリ、
5, 15, 26, 46, 66, 86, 126…語彙作成部、
6, 17, 28, 48, 68, 88, 128…語彙記憶部、
16, 27, 47, 67, 87, 127…発声確率計算部、
21, 41, 61, 81, 121…音声認識用辞書作成装置、
22, 42, 62, 82, 106, 122, 141…音声認識装置、

29, 49, 69, 89, 129, 142…音響分析部、
30, 50, 70, 90, 130, 143…尤度計算部、
31, 51, 71, 91, 131, 144…音響モデル格納部、
32, 52, 72, 92, 132, 145…照合部、
53…取り込み部、
54, 74, 94…文字列情報抽出部、
55, 75, 95…抽出条件格納部、
73…ウェブページ情報取り込み部、
76, 96, 117…制御部、
77…ウェブページ表示部、
93…テレビ番組情報取り込み部、
97…テレビ表示部、
98…録画部、
99…再生部、
101…音声リモコン、
102…マイク、
103…スピーカー、
104…リモコン制御部、
105, 107, 113…通信回線、
108…情報家電機器、
109…機器制御部、
110, 115…メイン情報メモリ、
111, 116…サブ情報メモリ、
112…メイン情報出力部、
114…情報提供センター、
133…類似度計算部、
146…第1語彙記憶部、
147…第2語彙記憶部、
148…検索部、
149…選択部。

【図1】



【図4】

(1-1)	浦和/市/大久保	うらわ/し/おおくぼ
(2-1)	浦和/市大/久保	うらわ/しだい/くぼ
(2-2)	浦和/市大/久保	うらわ/いちだい/くぼ

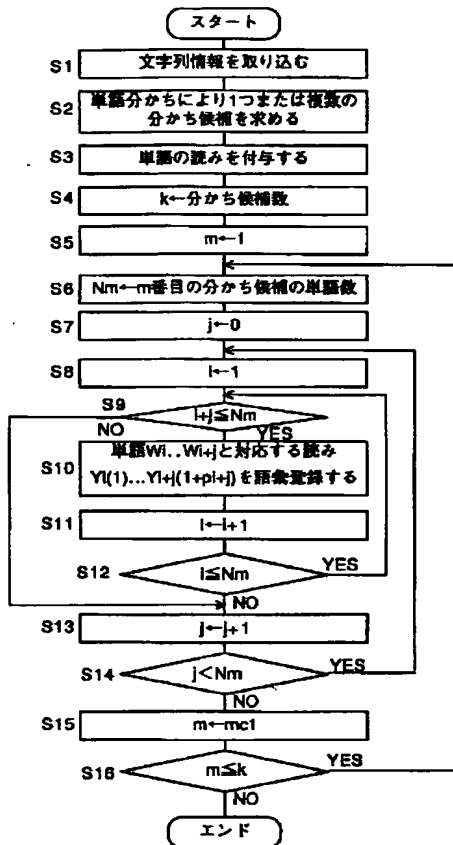
【図3】

(1)	浦和/市/大久保
(2)	浦和/市大/久保

【図5】

(1-1-1)	浦和	うらわ
(1-1-2)	市	し
(1-1-3)	大久保	おおくぼ
(1-1-4)	浦和市	うらわし
(1-1-5)	市大久保	しおおくぼ
(1-1-6)	浦和市大久保	うらわしおおくぼ
(2-1-1)	浦和	うらわ
(2-1-2)	市大	しだい
(2-1-3)	久保	くぼ
(2-1-4)	浦和市大	うらわしだい
(2-1-5)	市大久保	しだいくぼ
(2-1-6)	浦和市大久保	うらわしだいくぼ
(2-2-2)	市大	いちだい
(2-2-4)	浦和市大	うらわいちだい
(2-2-5)	市大久保	いちだいくぼ
(2-2-6)	浦和市大久保	うらわいちだいくぼ

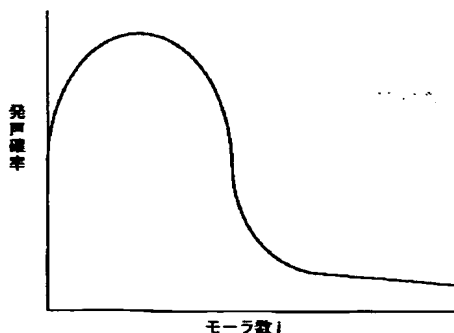
【図2】



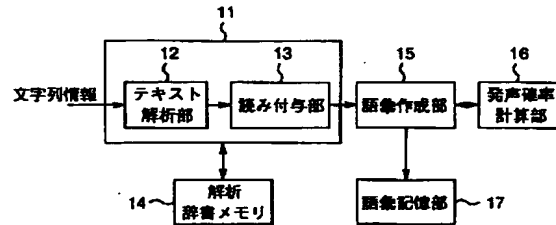
【図7】

(1)	浦和/市/大久保	KS(1)
(2)	浦和/市大/久保	KS(2)

【図10】



【図6】



【図9】

(1-1-1)	浦和	うらわ	WS(1,1,1)
(1-1-2)	市	し	WS(1,1,2)
(1-1-3)	大久保	おおくぼ	WS(1,1,3)
(1-1-4)	浦和市	うらわし	WS(1,1,4)
(1-1-5)	市大久保	しおおくぼ	WS(1,1,5)
(1-1-6)	浦和市大久保	うらわしおおくぼ	WS(1,1,6)
(2-1-1)	浦和	うらわ	WS(2,1,1)
(2-1-2)	市大	しだい	WS(2,1,2)
(2-1-3)	久保	くぼ	WS(2,1,3)
(2-1-4)	浦和市大	うらわしだい	WS(2,1,4)
(2-1-5)	市大久保	しだいくぼ	WS(2,1,5)
(2-1-6)	浦和市大久保	うらわしだいくぼ	WS(2,1,6)
(2-2-2)	市大	いちだい	WS(2,2,2)
(2-2-4)	浦和市大	うらわいちだい	WS(2,2,4)
(2-2-5)	市大久保	いちだいくぼ	WS(2,2,5)
(2-2-6)	浦和市大久保	うらわいちだいくぼ	WS(2,2,6)

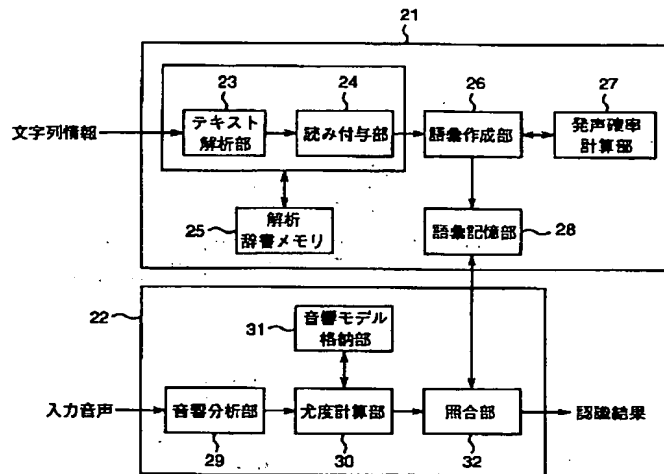
【図8】

(1-1)	浦和/市/大久保	うらわ/し/おおくぼ	YS(1)
(2-1)	浦和/市大/久保	うらわ/しだい/くぼ	YS(2)
(2-2)	浦和/市大/久保	うらわ/いちだい/くぼ	YS(3)

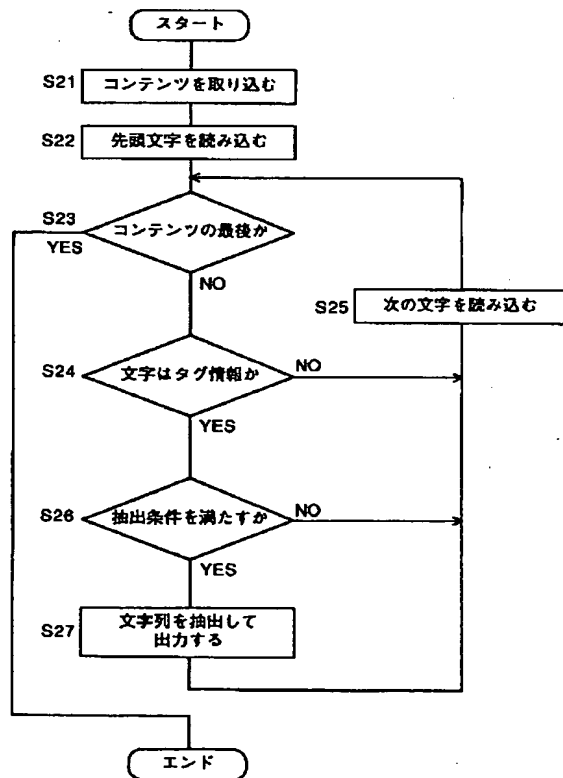
【図18】

日:	5/5
時間:	5:00 ~ 8:15
チャンネル:	XXX総合
ジャンル:	国内ニュース
G-code:	XXXXXXXX
番組名:	XXXニュースおはよう日本
番組内容:	ニュース、スポーツ、天気予報
出演者:	XXXXXXXX

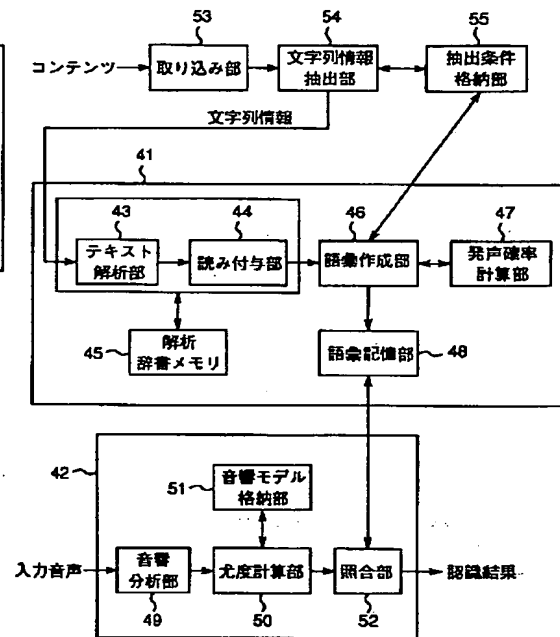
【図11】



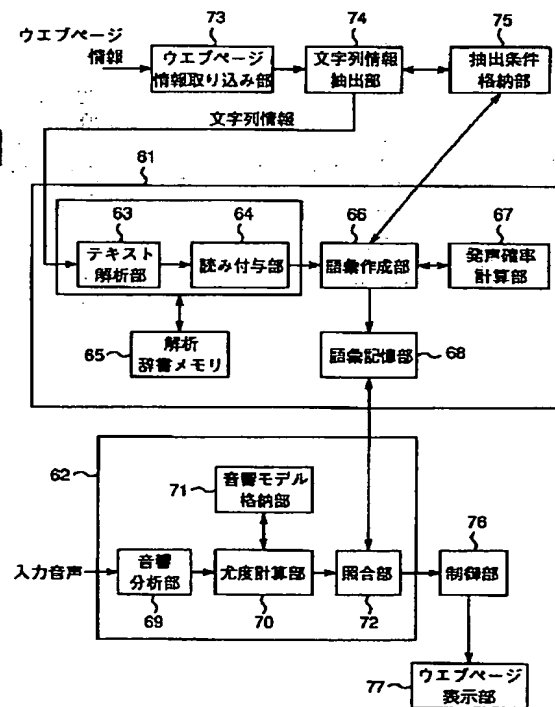
【図13】



【図12】



【図14】



【図15】

```

<html>

<head>
<title>
鈴木一郎のホームページ
</title>

<body>

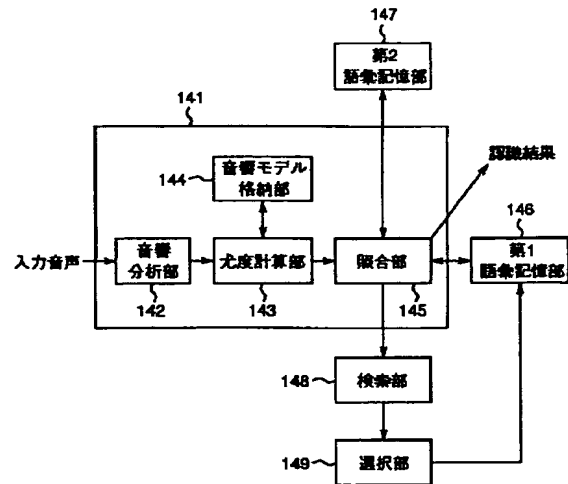
<h2> こんにちは </h2>
私は鈴木一郎です。<br>
ようこそ私のホームページへ

<br>
<a href="Intro.html">自己紹介します。
<br>
<a href="Intro.html#hobby">趣味のページ。
<br>
<a href="diary.html">日記です。
<br>

<address>
<a href="mailto:suzuki@mail.xxx.jp">suzuki@mail.xxx.jp</a>
</address>
</body>
</html>

```

【図22】



【図16】

68

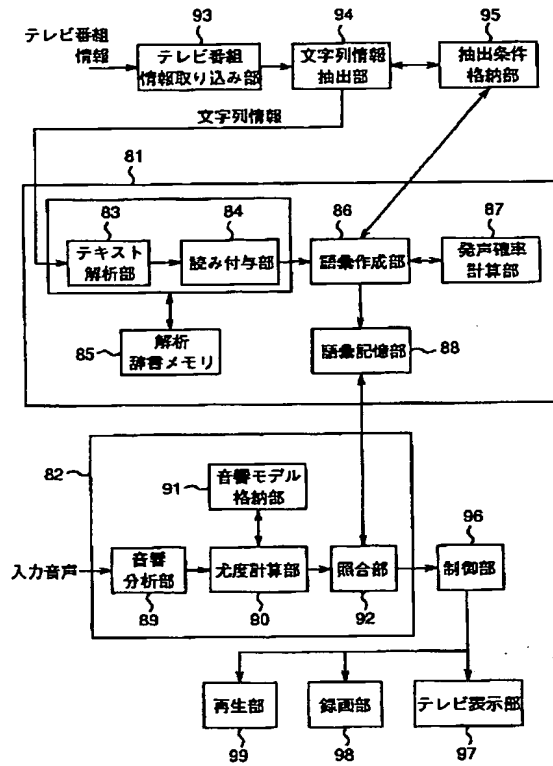
表記	読み方	発声確率	URL
鈴木	suzuki	P(1)	http://www.suzuki.xxx.jp
一郎	ichirou	P(2)	http://www.suzuki.xxx.jp
鈴木一郎	suzukiichirou	P(3)	http://www.suzuki.xxx.jp
鈴木一郎のホームページ	suzukiichirounoho-mupei-jl	P(4)	http://www.suzuki.xxx.jp
一郎のホームページ	ichirounoho-mupei-jl	P(5)	http://www.suzuki.xxx.jp
山田	yamada	P(6)	http://www.yamada.xxx.jp
太郎	tarou	P(7)	http://www.yamada.xxx.jp
山田太郎	yamadatarou	P(8)	http://www.yamada.xxx.jp
山田太郎のホームページ	yamadatarounoho-mupei-jl	P(9)	http://www.yamada.xxx.jp
太郎のホームページ	tarounoho-mupei-jl	P(10)	http://www.yamada.xxx.jp

【図19】

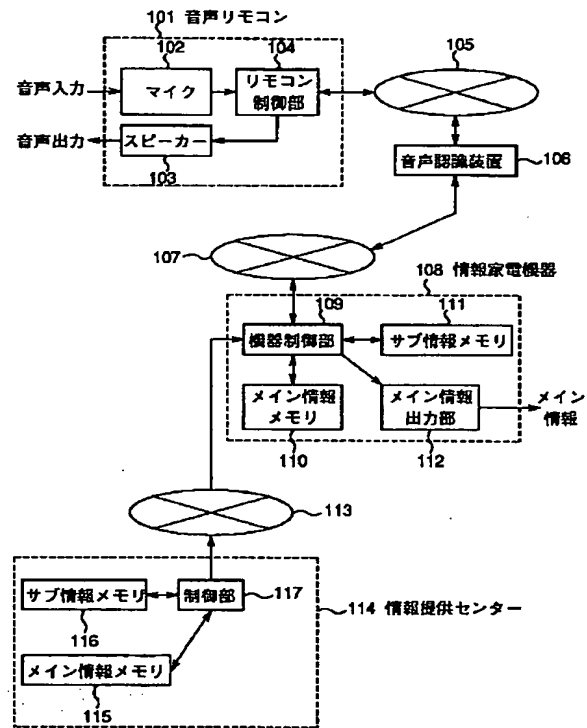
88

表記	読み方	発声確率	チャンネル	日	時間
XXX	enuechika-	P(1)	XXX放送	5/5	5:00 ~ 8:15
ニュース	nyu-su	P(2)	XXX放送	5/5	5:00 ~ 8:15
XXXニュース	enuechika-nyu-su	P(3)	XXX放送	5/5	5:00 ~ 8:15
おはよう日本	ohayounipon	P(4)	XXX放送	5/5	5:00 ~ 8:15
XXXニュースおはよう	enuechika-nyu-suohayou	P(5)	XXX放送	5/5	5:00 ~ 8:15
XXXX	mishio-moh	P(6)	XXX	6/6	20:00 ~ 21:00
サスペンス	sasupensu	P(7)	XXXテレビ	5/5	21:00 ~ 23:00
サスペンス制作劇場	sasupensu-seisakugakujokyo	P(8)	XXXテレビ	5/5	21:00 ~ 23:00
制作劇場	seisakugakujokyo	P(9)	XXXテレビ	6/6	21:00 ~ 23:00

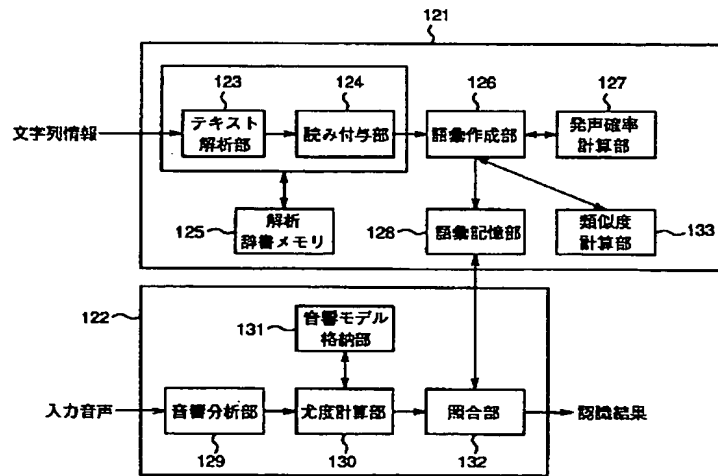
【図17】



【図20】



【図21】



This Page Blank (uspto)